DOCUMENT RESUME

ED 194 315 SE 033 007

AUTHOR	Goodrum, Denis, Ed.; Power, Colin, Ed.
TITLE	Research in Science Education, Volume 9. Proceedings of the Annual Conference of the Australian Science Education Research Association (10th, University of Western Australia, Perth, May, 1979).
INSTITUTION	Australian Science Education Research Association.
PUB DATE	May 79
NOTE	182p.; For Volume 8, see ED 180 768. Not available in hard copy due to copyright restrictions.
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Cognitive Processes; Cognitive Style; Elementary Secondary Education; Higher Education; Learning Problems; Science Curriculum; *Science Education; State of the Art Reviews
IDENTIFIERS	*Australia; *Science Education Research

ABSTRACT

This publication contains some of the papers and summaries of papers presented at the tenth conference of the Australian Science Education Research Association. Heading this collection are two invited papers by Peter J. Fensham and Colin Power which reflect the problems that confront science education researchers and which contemplate the future directions of this research. The remaining 21 papers cover a wide range of research interests and areas, covering general aspects of attitudes, learning difficulties with science concepts, curriculum issues, learning strategies, and thinking processes. (Author/CS)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Colin Power

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED194315

SE 033007

2

RESEARCH

IN

SCIENCE EDUCATION

VOLUME 9

Proceedings of the Tenth Annual Conference of the Australian Science Education Research Association, University of Western Australia, Perth, May, 1979.

EDITED BY: Denis Goodrum, Churchlands C.A.E.

GENERAL EDITOR: Colin Power, Flinders University

EDITORIAL BOARD: John Dekkers, Western Australian Institute of Technology
Arthur Lucas, Flinders University
Max Maddock, University of Newcastle
Michael Macfarlane, Darling Downs College of Advanced Education
Dick White, Monash University

BUSINESS MANAGER: Jim Butler, University of Queensland

AUSTRALIAN SCIENCE EDUCATION RESEARCH ASSOCIATION

All general correspondence concerning this publication should be addressed to Professor Colin Power, School of Education, Flinders University, Bedford Park, 5042, Australia.

Library Subscriptions to the Business Manager, Dr. J. Butler, Education Department, University of Queensland, St. Lucia, 4067, Australia.

Copyright 1979 by Australian Science Research Association.
Published by Australian Science Research Association.

ISSN 0157-244X

CONTENTS

		Page
Preface		(v)
Peter J. Fensham	Science Education Research: Present and Future	1
Colin Power	The Republic of Science Education Research	5
Leonie Rennie	Bloom's Theory of School Learning - A summary of the Implications for Science Education	13
R. Lewis and J.P. Keeves	Teaching Behaviour and Science Achievement	15
Alan R. Barton	A New Statistical Procedure for the Analysis of Hierarchy Validation Data	23
J.R. Northfield and L.C. Ingvarson	School-Based Developmental Research - From Ideas to Implementation	33
Peter J. Fensham	The Forest and the Woodchips - Are They Alternative Paradigms for Studying the Tree of Science Education?	43
M.R. Thorley, D.J. Boud, J.G. Dunn and T. Kennedy	The Aims of Science Courses	53
W.D. Ponniah	Students' Attitudes to an Inner-City Environment	55
M.N. Maddock	The Relationship Between Attitudes and School Achievement: A Further Exploration	69
Colin Power	Science Achievement and Attitudes at Primary and Secondary Interface	75
Roger J. Osborne and John K. Gilbert	Investigating Student Understanding of Basic Physics Concepts Using an Interview-About-Instances Technique	85
M.C. Kellett	Working Memory Capacity and Learning Difficulties	95

Contents - contd.

		Page
R.D. Linke and M.I. Venz	Misconceptions in Physical Science Among Non-Science Background Students: II	103
Darrell L. Fisher	The Impact of the Inclusion of ASEP Materials on Some Cognitive Outcomes in Different Types of Tasmanian Schools	111
Rod Fawns	The Treatment of Aspects of Change Associated with the Concept of Evolution in the Australian Science Education Project - The Development of an Argument Through Four Approaches to Content Analysis	119
P.P. Lynch	Scientific Language and the Tasmanian High School Child. The Tendency to Generalise	123
Gaye Rosen	The Effectiveness of the Use of Games as a Revision Technique in Junior High School Science	133
Abdul G. Khan	Audio-Tutorial Approach - A Multi-Media or an Individualized Method of Instruction?	144
John Dekkers and Anne Dowd	Science Classroom Practices and Design Characteristics of Instructional Materials: A Case Study	149
M.J. Sullivan and C.J. Dawson	Critical Thinking Ability in Science - What Can be Measured?	159
J.A. Rowell and C.J. Dawsor	Cognitive Conflict: Its Nature and Use in the Teaching of Science	169
Denis Goodrum	Creative and Logical Thinking in Adolescents	177
Contributors		185

PREFACE

This publication contains some of the papers and summaries of papers presented at the tenth conference of the Australian Science Education Research Association.

Since this conference marked the end of the organisation's first decade, it seemed appropriate to reflect on the problems that confront science education researchers in this country as well as contemplate the directions of where we should be going. To this end, Professors Fensham and Power were invited to present papers on these issues. These two papers head the list of enclosed papers and were the basis of a fruitful discussion session.

Although the discussions traversed many aspects of science education research, two common threads emerged. The first aspect concerned the problem of communication between the researcher and the teacher. There was concern that research findings were not filtering through to the classroom. The other area involved the kind of science education research being carried out. The need was expressed to research areas which were perceived as being more relevant to the classroom. Perhaps this is the challenge of the eighties: to initiate more useful classroom oriented research as well as to more effectively inform teachers of it.

The remaining papers cover a wide range of interests and areas. The initial group of papers are involved with more general aspects of science education. There are then three papers on different aspects of attitudes followed by papers on learning difficulties with science concepts. Next there are a number of papers examining curriculum issues and learning strategies. The final papers discuss research about thinking processes.

Colin Power
General Editor

Denis Goodrum
Editor

Research in Science Education,1979,9,1-4

SCIENCE EDUCATION RESEARCH: PRESENT AND FUTURE

Peter J. Fensham

Reflecting on growth areas in science education research that stood out as particularly noteworthy when I was on research leave in 1978, I find two umbrella areas to discuss.

The first field is concerned with cognitive structure and development in relation to scientific knowledge and skills. Piagetian research is now very lively in the U.S.A. This is somewhat surprising since ideas were less overtly acknowledged there in the big science curriculum era than they were in Australia or Britain. A.S.E.P. as we all know, was a science curriculum project directly related to Piaget's ideas. It has been rather tantalisingly kept from ready access to U.S. and other science educators beyond our shores, but is quite well known among some of the researchers.

In Australia, Tisher, Dale and Blake are, of course, well known in ASERA for their attempts to extend Piagetian identifications to groups of students, and Rowell and Calwell in Adelaide have contributed important theoretical possibilities which will be mentioned again below. There is certainly much continued interest in Piagetian ideas of concept development and their implications for curriculum development and teaching. Shayer (1978) in England is also still vigorous in this field and is now relating the data on students to teaching strategies. The Neo-Piagetian work is, however, I believe, of greater interest and here the work of Easley (1978) in the U.S.A. and his former colleague, Driver, at Leeds, is perhaps the most significant. They have found the Piagetian criteria for identifying stages of development restricting and too rigid and are attempting to develop more flexible criteria for assessing the responses of students to questions about scientific concepts and problems.

Although their interest was usually derived from sources other than Piaget, a number of other research workers are showing a growing convergence re problems and methods of investigation with these Neo-Piagetians. The general problem area is the nature and evolution of the cognitive structure of learners and the methodology increasingly involves clinical interviewing.

The early work of Shavelson using word association methods is still of interest and Preece (1978) in Exeter has made it better known in Britain. However, a number of other approaches are now very much in evidence, and it was quite striking how many workers from such varied backgrounds are now finding that clinical interviewing is an essential procedure to get at the all important prior knowledge or the ordering of the knowledge acquired in science learning. Anti-Piagetians of various ilks like Novak (1978) and others at Cornell and White and Gagné (1978) are now asking students to articulate their mental processes as they are confronted with scientific situations and problems. Pines of Cornell (now Maine) took up the complex

task of how to analyse and categorize the extended verbal data that emerges, and he has recently developed a useful guide to these procedures for research. White and several co-workers at Monash are attempting to relate eidetic and episodic forms of memory to the cognitive structure students derive from those classic aspects of science education - classroom demonstrations and practical work.

Rowell and Calwell (1978) have excited a number of Europeans with their suggestions about the similarity of the movement of students through Piaget's cognitive stages and the evolution of scientific concept and theories historically. Lybeck (1978) at Goteburg and Brumby at Surrey have, I believe, important empirical findings in the learning of physics and biology that confirm this suggestion and that point to an occupational rigidity among teachers about what is "correct" learning in science. Science teachers are, it seems, very likely themselves to confuse the models of their subject with the realities of phenomena and to be, as the "middle men", far less sensitive to the experience of scientific phenomena than are their research mentors or their students.

The second field of current research interest is the Contextual Influences on Science Education. Only after nearly two decades of incredible activity in the development of new science curricula, are a number of researchers becoming aware that science education is not a thing completely in its own right but is merely part of a much wider curriculum which, in its turn, is merely one part of schoolings' institutional response to the demands of the wider society in which it is set. Bowles and Gintis (1974) rightly pointed out that the curriculum developers (and I might add we science education researchers) operated as if the teaching and learning of science was a phenomenon that was insulated by a social and political vacuum. The recent state of the ark study of science education in the U.S.A. by Stake and others (1978) at Illinois has clearly spelt out how wrong we were. In Europe Layton (1973), Lundgren (1977) and Bourdieu (1977) have been asking us for some time to look outwards rather than inwards to explain or understand the inevitable shortfalls that have accompanied all the new science curricula. There is now much interest in this broader canvas and a number of authors at the Bat-Sheva Seminar (1978) [Shulman and Atkin (U.S.A.), Connelly (Canada), Eden (Israel), and myself] drew attention to the need to amend our traditional research designs and techniques if these new factors are to be considered adequately. Stake and Hamilton (1978) have used case study methods while Lundgren combined psychometric and more open procedures. Layton and Walker (1974) use historical methods and we have found these to be useful in some studies at Monash.

The contextual influences tend to be viewed in "political" terms (Atkin, Lundgren, Layton, Fensham) or in "cultural" ones (Stake, Connelly, Walker) but in either view, research methods which amplify and describe rather than narrow and count seem to be needed.

There are, of course, a number of other new fields of interest such as the work of the APU group in science education, Leeds, Chelsea, and Mirriam Perretz' work in Israel and U.S.A. on "curriculum potential." However, in general, I found little to excite me in the groups pursuing more traditional research approaches. Moreover, two large new fields of research to follow and perhaps, in which to contribute, are more than

challenge enough.

REFERENCES

ATKIN, M. Politics of curriculum innovation, development and implementation. In Bat Sheva Seminar Proceedings, 1978.

BAT SHEVA SEMINAR. Curriculum implementation and its relationship for curriculum development in science. A report on the proceedings. Science Education Centre, Hebrew University, Israel, 1978.

BEN PERRETZ, M. The concept of curriculum potential. Curriculum Theory Network, 1975, 5(2), 151-159.

BOURDIEU, P. & PASSERON, J.C. Reproduction in education, society and culture. London: Sage, 1977.

BOWLES, S. & GINTIS, H. Schooling in capitalist America. New York: Bonic Books, 1976.

CONNELLY, M. Implementation evaluation and professional development of science teaching. In Bata Sheva Seminar Proceedings, 1978.

EASLEY, R. & DRIVER, J. Pupils and paradigms, Studies in Science Education, 1978, 5, 61-84.

EDEN, S. Curriculum implementation as a process of innovation. In Bat Sheva Seminar Proceedings, 1978.

FENSHAM, P.J. Control and freedom in Australian secondary science education. In Bat Sheva Seminar Proceedings, 1978.

HAMILTON, D. Handling innovation in the classroom. In Reid, W.A. & Walker, D.F. Case studies in curriculum choice. London: Routledge & Kegan Paul, 1976.

LAYTON, D. Science for the people. London: Allen & Unwin, 1973.

LUNDGREN, U. Model analysis of pedagogical processes. Department of Educational Research, Stockholm Institute of Education, 1977.

LYBECK, L. A research approach to science education at Goteburg. European Journal of Science Education, 1978, 1, 119-124.

NOVAK, J. An alternative to Piagetian psychology for science and mathematics education. Studies in Science Education, 1978, 5, 1-30.

PINES, A.L. Scientific concept learning in children. Ph.D. thesis, Cornell University, 1977.

PREECE, P.F.W. Exploration of semantic space. Science Education, 1978, 62, 547-562.

ROWELL, J.A. & CAWTHORN, E.R. Epistemology and science education. Studies in Science Education, 1978, 5, 31-60.

SHULMAN, L. Implications of research on teaching for curriculum development and implementation. In Bat Sheva Seminar Proceedings, 1978.

STAKE, T. et al. Case studies in science education. See Atkin, M. In Bat Sheva Seminar Proceedings, 1978.

WALKER, R. Classroom research: A view from SAFARI. SAFARI-PROJECT, C.A.R.E., University of East Anglia, 1974.

WHITE, R.T. & GAGNÉ, R.M. Memory structures and learning outcomes, Review of Educational Research, 1978, 48, 187-222.

Research in Science Education,1979,9,5-12

THE REPUBLIC OF SCIENCE EDUCATION RESEARCH

Colin Power

INTRODUCTION

In May 1970, a small group of academics accepted Peter Fensham's invitation to attend the first conference on science education research to be held in Australia. That conference led to the decision to continue to meet informally. At the next meeting, held at Macquarie University, it was decided to establish an Australian Science Education Research Association. In the new republic, called ASERA, the supremacy of the people was guaranteed by the simple expedient of having no ruling elite. On this, the tenth anniversary of our first meeting, it is appropriate to ask where the republic is, and ought to be heading.

THE REPUBLIC OF SCIENCE EDUCATION RESEARCH

In reflecting on this question, I began to dream of an ideal state, a republic of education founded on truth, freedom, and equality. My vision was of a promised land in which the formulation of educational policy was rational, educational resources were distributed equally, and the supremacy of the people (teachers, parents, students and researchers) over the system held sacred. I envisioned a Platonic golden age. Consensus and absolute certainty had been reached about what is of value in science education by the philosophers. The experimentalists were able to provide certain knowledge how these ends could most effectively be taught and learned. A Republic of Science Education administered, but not controlled, by the just and the wise on behalf of all.

But we are far from this promised land. So I decided to consult some of the philosopher-kings at Mount Flinders to see if they could help me find the path out of the wilderness. But alas! They argued that there is no such thing as science education research. There is no promised land, no republic of science education. In their view, a claim to the territory of science education is about as supportable as Prince Leonard's proclamation of the Hutt River Province. With my vision evaporating, I turned to the problem of where science education research is heading.

SCIENCE EDUCATION RESEARCH: ITS SCOPE AND ITS LIMITS

Each area of research is concerned with some part of reality, the delineation of which defines its territory. The particular part of reality studied by science education research workers was defined at our first meeting:

> Science education is concerned with translating the content and models of inquiry of the scientific discipline into materials and sequences of learning experiences appropriate to the needs of students of varying abilities and interests. The process may be seen as involving many critical decisions about what should be taught, when, how and under what conditions... (Power, 1970).

However, as my colleagues at Flinders University reminded me, parts of this territory are claimed by philosophers, psychologists and sociologists, as well as science educators. Moreover, science education research workers almost invariably use the concepts, the methods, the research paradigms of the above disciplines in studying science education phenomena. For science education research to be recognized as a discipline, it would need to focus on a large enough body of inter-related phenomena, theories and problems to justify its isolation from other areas of research (Hirst, 1974). In this sense, science education research as it is practised today is not a discipline.

To the degree that disciplines like history, philosophy, psychology and sociology do have structures and well-developed research tools, they can contribute to the task of providing increasingly refined and elaborated maps of parts of the territory of science education. But in my view, some of the most significant problems of science education cannot and ought not be reduced to fit the puzzle form imposed by a discipline like psychology or sociology. As several contemporary educational researchers and philosophers (Crittenden, 1973; Lundgren, 1977; Power, 1977; Gowin, 1972) have argued, educational research ought to be more than a smorgasboard of applied disciplines. As Gowin (1972) puts it: "many events which are are educational never get studied, and many events with which educational researchers concern themselves now have little or nothing to do with education." There is room for new paradigms in educational research, for research in which education is not only the subject, but also the objective of the research.

Alongside the rigorous attempt to explain the psychological processes within the individual learner or teacher in the science classroom, there is a place for attempts to account for the pedagogical processes taking place in educational systems, schools and classrooms. Shulman (1974) hinted at this in his call for a renascence of a modern form of the psychology of school subjects. Shulman argues that the proper scope for education theory is what he calls "theories of the middle range" and that the the function of theory is to help teachers understand or render meaningful the educational phenomena of the contexts in which they work. In an earlier article, I (Power, 1976) argued that we need to start with an analysis of the procedures of science education so as to be able to identify the problems which are significant and which pass as being distinctively educational.

A few individuals have been working on the development of theories based upon reflections on the results of classroom research and educational practices, rather than perspectives borrowed from psychology or sociology. The attempt to create educational constructs and a propositional structure which accounts for regularities in distinctively educational phenomena is still rather feeble. One can point to constructs like the steering group, piloting (Lundgren, 1977), with-it-ness, smoothness (Kounin, 1970),

trade-off and balance (Good & Power, 1976), and to the attempts of Westbury (1973), Lundgren (1977), Good and Power (1976), and others to build pedagogical models.

The call then is for "pedagogical theories out of pedagogical practice," rather than "psychological theory into practice." It is towards the study of the distinctively educational problems of science education that I believe at least some part of science education research ought to be directed.

THE PURPOSES OF SCIENCE EDUCATION RESEARCH

What is the purpose of research in science education? What ought to be its purpose? Should we be attempting to explain the phenomena of science teaching, or be attempting to improve science teaching practice? The so-called conflict between theory and practice has been with us ever since the possibility of treating education as a science has been entertained.

Looking through the recently published analyses of the NARST survey of the priorities for research in science education (Butts, et al. 1978, Okey & Yeany, 1978; Yeany & Capie, 1978) we see that in the U.S., the move is away from research for the purpose of developing and testing of theories and towards research for the purpose of improving practice. There can be little doubt that the trend is a reaction to the alleged failure of educational research to contribute to practice, and that we will begin to feel the winds blowing in the U.S.

Against this tendency towards decision-oriented, applied, practical research, there are those, like Kerlinger (1979), who argue the case for "basic" research: the purpose is theory, understanding and explanation. Research has no other purpose. Kerlinger cites evidence which he claims indicates that basic research is likely to have a greater effect on practice than applied, programmatic or targeted research. He claims that the effects of basic research are indirect and are felt only over appreciable periods of time.

When one reflects upon the impact of the work of Piaget, Ausubel, Bruner, Gagné and Bloom on curriculum development and evaluation in science, it is clear that Kerlinger's claims have some substance. Major theories do influence our thinking - the problems we recognize, the alternatives we generate, the ways we evaluate these - and hence our practice. A research study is rarely important in itself, but only acquires importance if it fits into a system of findings or beliefs.

While I tend to take the conservative position that the major purpose of research is theory, understanding or explanation, I reject emphatically the proposition that one can and should separate the two main aims of research as if they were poles which pull against each other. Educational research (as opposed to research on education) for theory generation never can detach itself from the problems of practice; evaluation and case study research never can be devoid of theory, however unsystematic and submerged that theory may be. There is no logical or practical reason why research

conducted by those interested in solving practical problems cannot be scientific (Shaver, 1979).

To regard the pursuit of understanding irrespective of the problems of practice as being only for the intellectual satisfaction of research workers is not just an oversimplification, but a serious error of categorization. As Jevons (1973) has argued, the main functions of science are social ones. So too MacRae (1976) asserts, the function of social science is not simply to serve the interests of academics themselves, but:

> Rather, it is to provide guidance to society, through research, reasoned discourse, and education as to what interests should be served in particular circumstances, and as to the means to do so. (p. 306).

MacRae goes on to argue for the fusion of scientific research and ethics in the social sciences. What is desirable in practice cannot be decided on scientific grounds alone, reasoned ethical analysis is also essential. From this perspective the purpose of science education research is to help us understand the ethical and the technical problems of practice.

Because science teaching is a deliberate effort to promote the learning of beliefs, attitudes, skills and concepts thought to be of significant human value, one cannot engage in research in science education without having regard for what will improve it, or without serious consideration of the practical and ethical implications of any conclusions reached: otherwise it is not science education research.

In essence, I am arguing that the purpose of science education research ought to be understanding for the purpose of improving practice. The research ought to lead to theories of the "middle range" which are distinctively educational in that they address educational problems and facilitate the continuous education of science teachers.

THE CONTENT OF RESEARCH AND POLITICS WITHIN THE REPUBLIC

When we look through the issues of Research in Science Education, it becomes clear that, cumulatively, the research does have something to say about the development and implementation of curriculum materials (notably ASEP) and the problems of measuring student outcomes in science. In addition, one finds the odd paper which questions something taken for granted by most. But apart from analyses of curriculum materials and student misconceptions, little of what we have done really does seek to illuminate a problem which science teachers recognize as being significant. Moreover, the results and their educational implications are rarely communicated to teachers, and researchers rarely focus on the problems which concern science teachers.

But no matter what problems we believe ought to be studied, the problems actually studied in the future will be strongly influenced by the prevailing research paradigms. After all these paradigms define which areas of the territory of science education are worth mapping. As I have indicated in other papers (Power, 1976; 1977) the prevailing empirical paradigm is under attack. In the future, advocates of alternative

paradigms will become politically more active and more successful. In particular, the problems defined as significant by the "interpretative paradigm" will become the focus of attention. ASERA conferences will have their share of papers looking at the ways science teachers and students define their situation, and the common sense knowledge and theories which they use to guide their actions in the classroom.

In addition to the forces within the academic community itself, science education research will increasingly be determined by changes in the nature of the power relationships among educational administrators, teachers and researchers. Currently, we are witnessing a redistribution of power over educational matters. For example, we have seen greater autonomy and responsibility for curriculum and evaluation handed to the schools by the central authorities.

As a consequence of such shifts, most of the science education research done in schools will be research which the school and the science teachers want to be done. It will be client-oriented, and client-controlled, if not school-initiated (action research). Essentially, this influence will be a constructive one. However, some of the most significant educational problems are likely to be those which teachers and schools are either unconscious of or will want to keep hidden. For instance, I believe we ought to initiate research on what schools are doing about the incompetent and uncommited science teacher. But I suspect the politics of education are such that the problem science teacher will continue to be swept under the carpet. Both schools and researchers have a vested interest in maintaining the division of labour and maintaining a communication gap.

THE CONTENT OF RESEARCH AND POLITICS BEYOND THE REPUBLIC

As well as being determined by forces within the republic, the content of research is not impervious to external politics. The large scale expenditure on scientific research in the post-war period has led governments to plan the direction of scientific research. But freedom to choose one's own research problems and mode of attack is a cherished ideal. Only by allowing free play to the individual initiatives of researchers, argues Polanyi (1962), can we optimize the use of scientific resources. Yet the invisible hand of government directs much of the work of researchers to those areas which the advisors of governments see to be of greatest value, as well as to be the collection of data for the purpose of justifying government policy.

We can gain some insight into where the invisible hand of government agencies may lead us from an analysis of changes in funding arrangements for educational research. The National Institute of Education is committed to support research which will "improve educational quality and foster equality of educational opportunity." The Program's orientation is towards research strategies which recognize the values of:

1. bringing the perspectives of a variety of disciplines to bear on education related questions

2. encouraging collaboration between investigators and practioners in the defining of research questions and conducting research

3. conducting research in such a way that it benefits participants.

The 1979 Grants program extends research interests to the study of teaching learning processes in settings outside of school and throughout the life span. Areas given priority are literacy (cf. scientific literacy), mathematics (cf. mathematical tasks encountered in science lessons) and teaching (cf. science teaching in the home, community and work settings). Research in school settings given priority concentrates on organizational structure features of schools and classrooms (e.g. allocation of time; pacing and sequencing of instruction; management systems; incentives) and outcomes.

The ERDC has established four priority areas viz. Multicultural Education, The Impact of Demographic Effects and Social Change on Education, Recurrent Education, and What is Happening in the Classroom. The idea of devoting part of ERDC's funds to planned, sequential and related topics selected as priority areas has received wide support, at least from governments, community groups and state education authorities.

Areas assigned priority by the ERDC Advisory groups include: equality of school experience, ethnic identity, school leavers, adult learning techniques, recurrent education and life cycle needs; the needs of students in Grades 5-6 and 9-10, the expectations of students, parents and teachers, curriculum decisions and teaching strategies and decisions. The Schools Commission has indicated that it intends to focus on the school experience of 15-16 year olds and the transition from school to work.

The themes which emerge in the NARST, NIE, ERDC and Schools Commission priority statements reflect a concern for solving immediate social problems, and solving them quickly. Many of these problems are merely the surface manifestations of more serious problems which are deeply embedded in the economic, political and social structures of our society. The more clearly our research indicates the linkage between the existing patterns of distribution of wealth, power and prestige in our society and the ways in which these manifest themselves in educational systems, the more we will need to fight for the right to challenge existing practices and assumptions with evidence and logic. But our arguments will be heeded only if we are accountable, that is, our work is seen as having a significant and recognizable impact on the quality of science education.

METHODS OF RESEARCH: TOOLS AND TOIL IN THE REPUBLIC

The forces which influence what is to be mapped in science education also will influence who does the mapping, how and with what tools. It seems to me that we will no longer be able to map by bulldozing our way into new terrain, upsetting its fragile ecology and manipulating its inhabitants according to the dictates of a classical research design. In the new republic, the Kerlingers and the Campbell and Stanley's of science education will no longer toil in anything other than their own backyard.

Mass survey methods will be used in the manner of aerial surveys for the purpose of discovering hidden resources and spying on suspected trouble spots. But the mass descriptive survey and the more sophisticated of the explanatory naturalistic surveys will become more and more the preserve of governments and research syndicates. The difficulties encountered in convincing more than a few schools or teachers that what we propose will significantly benefit them will make the case study, the cooperative evaluation, and action research the norm for most toiling at science education research. Increasing use will also be made, therefore, of qualitative methodologies and unobstrusive measures. Initially, the qualitative researchers will begin with fumbling fingers and primitive tools to produce crude maps of specific science education cultures. In time, the principles of procedure and the canons of inquiry will themselves become the object of study, and the rules and tools more explicit. The tools which need to be developed are those which researchers and teachers can use in continuous monitoring of the consequences of educational decisions. The data analysis techniques will need to include procedures for detecting significant shifts in the operation of dynamic systems and ways of amalgamating case records and case studies of action research projects in ways which can inform decision-making and sensitize policy-makers to the realities of life in classrooms and schools.

CONCLUSION

In this paper I have tried to examine the possibility of a free and independent republic of science education in which all are committed to the pursuit of truth, freedom, quality and equality. The task will undoubtedly be a difficult one, given the constraints and the external forces which are likely to keep us in the wilderness. But the way to the promised land is, I believe, to develop individually and collectively, a commitment to questioning the justice as well as the effectiveness of what we are doing.

To count as science education research, the work we do must contribute to the education of teachers. It must help sensitize teachers to the nature of the problems which confront them. It must help them to make professional judgments about what is justifiable, feasible and worth trying, given what we know of the effects of different strategies. And it must contribute to the development and refinement of the common-sense knowledge and pedagogical theories by means of which science educators make sense of the phenomena of science education and which guide their actions.

REFERENCES

BUTTS, D. et al. Priorities for research in science education: A Delphi study. Journal of Research in Science Teaching, 1978, 15(2), 109-114.

CRITTENDEN, B. Values in the study of education. Paper presented at Australian Association for Research in Education Conference, 1973.

GOOD, T.L. & POWER, C.N. Designing successful classroom environments for different types of students. Journal of Curriculum Studies, 1976, 8(1), 45-60.

GOWIN, D.B. Is educational research distinctive? In L.G. Thomas (Ed.) Philosophical redirection of educational research. 71st NSSE Yearbook (Part 1) Chicago: University of Chicago Press, 1972.

HIRST, P. Knowledge and the curriculum. London: Routledge & Kegan Paul, 1974.

JEVONS, F.R. Science observed. London: Allen & Unwin, 1973.

KERLINGER, F. Behavioral research: a conceptual approach. New York: Holt, Rinehart & Winston, 1979.

KOUNIN, J.S. Discipline and group management in classrooms. New York: Holt, 1970.

LUNDGREN, U. Model analyses of pedagogical processes. Lund, Sweden: C.W.K. Gleerup, 1977.

MACRAE, D. The social function of social science. New Haven: Yale University Press, 1976.

OKEY, J.R. & YEANY, R.H. The relationships among science educators' research priorities and demographic variables. Journal of Research in Science Teaching, 1978, 15(5), 387-394.

POLANYI, M. The republic of science. Minerva, 1962, 1, 54-59.

POWER, C.N. A review of current research in science education in Australia. Australian Science Teachers' Journal, 1970, 16(2), 51-55.

POWER, C.N. Competing paradigms in science education research. Journal of Research in Science Teaching, 1976, 13(6), 575-587.

POWER, C.N. A critical review of science classroom interaction studies. Studies in Science Education, 1977, 4, 1-30.

SHAVER, J.P. The productivity of educational research and the applied basic research distinction. Educational Researcher, 1979, 8(1), 3-9.

SHULMAN, L.S. The psychology of school subjects: a premature obituary. Journal of Research in Science Teaching, 1974, 11(4), 319-339.

WESTBURY, I. Conventional classrooms, 'open' classrooms and the technology of teaching. Journal of Curriculum Studies, 1973, 5(2), 99-121.

YEANY, R.H. & CAPIE, W.R. A factor-analytical study of the research priorities of science educators. Journal of Research in Science Teaching, 1978, 15(2), 167-172.

Research in Science Education, 1979, 9, 13-14

BLOOM'S THEORY OF SCHOOL LEARNING—A SUMMARY OF THE IMPLICATIONS FOR SCIENCE EDUCATION

Leonie Rennie

This paper contends that Bloom's Theory of School Learning has significant implications for science educators. This theory, described in Bloom's recent publication Human Characteristics and School Learning, attempts to explain, to predict and to modify individual differences in school learning. It is based upon research connected with mastery learning strategies and takes the view that an effective learning system is characterised by high achievement together with minimal variation in student learning. The theory proposes that only three major variables, the cognitive and the affective entry characteristics of the learners, and the quality of instruction, account for the nature of the learning outcomes (both cognitive and affective) of each learning task. Bloom has argued his theory both logically and plausibly - certainly this is the consensus of reviewer opinion - and the purported relationships between variables are consistent with empirical evidence. In particular, the cognitive aspects of the theory are well established. Research in affective areas, while not extensive, is also supportive. The limited amount of research designed to test predictions made from the theory is also supportive, although Bloom's claim that variability in learning rates can become negligible remains controversial.

If it is accepted that Bloom's theory offers a useful framework for thinking about student learning, then its implications for science education concern both science teaching and science education research. The implications concerning science teaching recognize that students usually begin regular science instruction during their first high school year and already possess widely differing entry characteristics. Even so, the opportunity exists for teachers to become more aware of, and to make changes in, certain aspects of student learning. For example, Bloom's theory emphasises the importance of the early learning tasks, both to achieve mastery, which provides the necessary cognitive behaviours requisite to later learning tasks, and to achieve the success which will improve and maintain positive affective characteristics. The quality of instruction must be matched to individual learners, with extra time spent in early learning tasks serving as a preventative, rather than a remedial measure.

Recent publications have expressed a need for a more theoretical base in science education research. Many topics are poorly or inconclusively researched. The National Association for Research in Science Teaching has recently identified a number of priority research areas and Bloom's theory is potentially relevant to many of these. The theory can be useful in two ways. First, as an organizational framework, it can be used to bring among discrete pieces of research, and to provide researchers with a systematic approach to new research, based upon a conceptualisation of, and interrelationships among, the relevant variables. Second, the theory

stimulates further research by pointing to problem areas, suggests new ways for tackling these, and enables the use of a hypothetico-deductive design. To illustrate this capacity of Bloom's theory this paper concludes with an outline of the author's current research into affective aspects of science education. In keeping with Bloom's approach, affect related to science is defined as a resolution between the student's perceptions of his past history in science, his likely future success, and the usefulness to him of the present learning task. The research examines the reciprocal relationship between science-related affect and achievement in science, and tests certain qualitative and quantitative hypotheses deduced from Bloom's theoretical model. The findings will clarify the role teachers play in developing positive affect towards science and science learning.

REFERENCES

BLOOM, B.S. Human characteristics and school learning. New York: McGraw-Hill, 1976.

Research in Science Education, 1979, 9, 15-22

TEACHING BEHAVIOUR AND SCIENCE ACHIEVEMENT

R. Lewis and J. P. Keeves

INTRODUCTION

In 1971 Keeves reported the findings of an extensive study of the relationships between classroom processes and curriculum outcomes in science and mathematics at the Form 1 level in Canberra (Keeves, 1971). The sample of the study was almost the entire Form 1 population and involved 15 schools and 72 classes. This secondary analysis of the data is based on a little more than 80 per cent of the population.

METHOD

Independent Variables.

In the original study, data were gathered on 146 variables measuring student characteristics and performance, teacher background, classroom practices and structural properties of the classroom. The variables that are pertinent to this paper are the 32 observed teacher behaviours. These comprise the Classroom Observation Schedule.

The design and development of the observation schedule was based on the discussion of the measurement of classroom behaviour through systematic observation by Medley & Mitzel (1963) and Anthony (1967). The schedule was designed for observation of 40 minute class periods. The total amount of time of observations was approximately 160 minutes. The basis for deciding upon four visits of 40 minutes was primarily the work of Armstrong, Vere De Vault and Larsen (1967). However, it is interesting to note the recent work of Rowley (1978) which indicates that four of five visits totalling approximately 150 minutes of observation can result in reliability coefficients of between 0.65 and 0.70 and that little additional reliability may be obtained by increasing either the length of, or number of visits.

The observational variables employed in this study are described in summary form in Table 1. A fuller description of these and of the Classroom Observation Schedule have been published previously (Keeves, 1974).

Although it would be possible, in retrospect, to criticize the selection of variables because some are clearly of little current interest, at the time of planning the original study the selection was considered appropriate, since the variables were seen as reflecting important dimensions of the classroom environment.

TABLE 1 Item Statistics, Correlation and Regression Coefficients for Science Process Characteristics

	Notes	Between Student		Between Classroom		Between Classrooms Within School		Between School	
		r	β	r	β	r	β	r	β
Time Used by Teacher	Time in minutes	0.07	0.04	0.14	0.02	0.19	-0.01	0.06	0.23
Students Listen	Time in minutes	0.16	0.08	0.28	0.07	0.48	0.08	0.06	0.19
Students Talk	Time in minutes	-0.01	-0.00	-0.05	0.02	-0.25	0.04	0.24	0.04
Question and Answer	Time in minutes	0.14	0.07	0.21	0.06	-0.22	-0.04	0.46	0.25
Students Write	Time in minutes	-0.23	-0.10	-0.34	-0.07	-0.57	-0.08	-0.40	-0.22
Students Read	Time in minutes	-0.07	-0.04	-0.03	-0.03	0.16	-0.05	0.30	0.11
Students Investigate	Time in minutes	0.03	0.01	0.04	-0.01	0.53	0.11	0.05	-0.08
Students Mark Work	Time in minutes	-0.01	-0.00	-0.08	0.00	0.15	-0.01	-0.22	0.00
Unclassified	Time in minutes	-0.11	-0.08	-0.17	-0.08	0.13	0.09	-0.29	-0.33
Change in Activity	Number of changes	0.00	0.03	0.04	0.04	0.11	0.01	0.18	0.23
Teacher Reviews Work	Time in minutes	-0.01	0.03	-0.02	0.04	-0.58	0.01	-0.22	0.02
Teacher student work contacts	Number of instances	-0.27	-0.11	-0.51	-0.05	-0.39	0.09	-0.48	-0.19
Teacher Asks Question	Number of instances	0.03	0.03	0.03	0.03	-0.23	-0.09	0.13	0.12
Student Asks Question	Number of instances	0.15	0.03	0.26	-0.02	0.22	-0.00	0.62	0.13
Invitation to Participate	Number of instances	0.04	0.00	0.09	-0.01	0.21	-0.08	0.02	-0.10
Use of Languages	Number of instances	-0.08	-0.02	-0.06	0.01	0.03	-0.02	-0.18	0.07
Invitation to Inquire	Number of instances	0.03	-0.01	0.09	-0.03	0.44	-0.04	0.37	0.05
Consider Work Habits	Number of instances	-0.07	-0.01	-0.14	0.01	0.31	0.09	-0.36	-0.05
Raise Aspirations	Number of instances	0.04	-0.00	-0.05	-0.02	-0.03	-0.01	0.22	-0.01
Casual Praise	Number of instances	-0.07	-0.02	-0.15	0.01	-0.40	0.00	-0.04	0.07
Deliberate Praise	Number of instances	-0.00	0.03	-0.05	0.07	-0.25	-0.05	-0.29	0.16
Other Reward	Number of instances	-0.20	-0.07	-0.31	-0.02	-0.14	-0.05	-0.34	-0.07
Total Praise	Number of instances	-0.08	-0.02	-0.16	0.02	-0.40	-0.01	-0.10	0.08
Casual Rebuke	Number of instances	-0.20	-0.07	-0.41	0.00	-0.23	0.06	-0.45	-0.11
Deliberate Rebuke	Number of instances	-0.17	-0.06	-0.29	0.00	-0.56	0.13	-0.00	0.00
Other Punishment	Number of instances	-0.15	-0.06	-0.27	-0.01	-0.56	-0.04	-0.18	-0.06
Total Rebuke	Number of instances	-0.22	-0.08	-0.41	-0.00	-0.57	0.11	-0.22	-0.05
Positive Support	Number of instances	0.02	0.06	-0.12	0.10	-0.19	0.13	-0.17	0.15
Negative Support	Number of instances	0.16	0.06	0.22	0.02	-0.03	-0.09	0.44	0.14
Laughter With	Number of instances	0.13	0.06	0.21	0.03	0.43	-0.00	0.26	0.11
Laughter At	Number of instances	0.19	0.09	0.26	0.06	0.13	-0.03	0.08	0.13
Autonomy	Number of instances	0.04	0.03	0.09	0.04	0.20	0.05	0.29	0.22

Note: Correlation and regression coefficients are underlined where:

Between student $r > 0.10$, $\beta > 0.10$ Between classroom within school $r > 0.24$, $\beta > 0.10$
Between classroom $r > 0.24$, $\beta > 0.10$ Between schools $r > 0.52$, $\beta > 0.10$

Dependent Variables.

Measures of science achievement were made at the beginning and end of the year. Both the pre and post-tests comprised selections of items from eight science content areas and two levels of cognitive functioning. The internal consistency measures (K.R.20) for the pre and post-tests were 0.78 and 0.84 respectively.

Reliability of Observation.

To establish the reliability of the observations, the source of error associated with differences between the two observers was first considered. For a sample of 10 classrooms, which were assessed by both observers simultaneously, both interrater and intraclass correlations coefficients were calculated to assess the reliability of the measures. Of the 32 observational measures, Keeves argued that the low reliability of seven placed doubt on their usefulness. These were behaviours numbered 3, 6, 10, 11, 12, 19 and 32 in Table 1.

A further consideration in determining the usefulness of the observational measures was the stability of the level of teacher behaviour. The importance of this was for purposes of generalization from the classroom behaviour which had been observed to behaviour on other occasions. To test the stability of his measures, Keeves considered the stability of observations between visits and between observers employed in the study. The results of these analyses indicated that the stability of 16 of the teacher behaviour measures were uncertain, largely due to the homogeneity of teacher behaviour reducing the variance between teachers.

Relationships Between Teacher Behaviour and Student Achievement.

The analysis relating teacher behaviour and student achievement were carried out using regression analysis. In a naturalistic study, where there has been no random assignment to a treatment group of either individual students or intact classrooms of students, it is impossible to argue that pretest performance does not influence the behaviour of the teacher and the students within the classroom group. Therefore analysis of covariance was considered to be inappropriate. Regression analysis partials out the effect of pretest performance from both teacher behaviour and post-test score. In addition, regression analysis provides readily understood results. Using beta coefficients, it is possible to estimate the change in criterion score which is related to a standard unit variation in the predictor variable.

Both zero order correlations and beta coefficients associating teacher behaviour and final science achievement with initial science achievement partialled out, are presented in Table 1 for each of the 32 predictors.

Level of Analysis.

Issues associated with the appropriate level and unit of analysis must be considered with respect to factors of design, namely the sampling procedures employed and the manner in which the treatment or learning conditions operated. There are, however, other issues relevant to determining the appropriate level and unit of analysis that are related to the research questions being investigated.

It is common both in naturalistic and experimental studies of teacher effectiveness to select a sample of schools for investigation. It is also common to proceed to a second stage of sampling and to select classrooms and teachers from within schools for testing and observation. However, it is important to note that the teachers selected from one school, with its unique 'tone', may not be statistically independent of one another and although they have had an equal probability of selection they may not form a simple random sample for the purposes of analysis. Consequently, in the analysis of the data it is essential to consider both the primary sampling units of schools as well as classrooms and classrooms nested within schools. However, once again, to provide comparisons with previous research, the analyses must also be conducted at a level of the individual student, and under some circumstances, students nested within classrooms.

When performing analyses at these five different levels appropriate aggregation and disaggregation of data must occur. At the classroom level, for example, class mean achievement in science must be used as the measure of the dependent variable. When considering the relative effectiveness of teachers within schools, school means need to be calculated for the teacher behaviour and class mean achievement data. It is then necessary to create deviation scores by subtracting the school mean from each teacher and class mean score. The further level of analysis with students overall requires each student to be given the level of teacher behaviour attributed to his class, that is the teacher scores are disaggregated. The analysis then proceeds as if the teacher scores are properties of the student.

Level of Significance.

The question of level of significance for both the zero order correlation coefficients and the standardized partial regression coefficients warrants consideration. The values are equivalent to those which would reflect a five per cent level of significance in generalizing to a population. However, it must be recognized that almost a complete population was under survey and generalization from the sample to the population is somewhat redundant. It should also be noted that Ross (1976) estimated the correlation coefficient design effects for students in intact classrooms to be approximately 5, and therefore this value has been employed to estimate the general level of significance for the between student correlations. Since no information was available on the magnitudes of the design effects for the other levels of analysis, no correlation for clustering has been made.

Ross (1976) has shown that the design effects for standardized regression coefficients are only slightly less than those for correlation coefficients and under these circumstances approximately the same values as employed for correlation coefficients have been used for examining regression coefficients. In the case of between student regression

coefficients the value of 0.10 would appear appropriate. For between classrooms, and between classrooms within schools the corresponding value should be 0.24.

However, this would appear very stringent with respect to previous work since approximately five per cent of the variance is explained by standardized regression coefficients of the order of 0.24. Therefore, a decision was made to consider those standardized regression coefficients which resulted in an explanation of one per cent of variance or more. It would appear that explained variance of this magnitude is of interest because many studies concerned with teacher behaviour report findings associated with variance of approximately this amount. The educational and statistical significance of such findings must however continue to be questioned.

RESULTS

The results of the analyses conducted at the four levels of between student, between classroom, between classroom within school and between school have been recorded in Table 1.

Between Student Analysis.

Analyses at this level allow for investigation of relationships between individual students and the classroom teacher. The use of individuals' scores on the science tests implies that systematic interactions between pupils and teacher will result in differential student gains in science achievement. However the findings recorded indicate that the results of the between student analyses very closely reflect those obtained at the between classroom level of analysis and therefore description and interpretation will be provided in the next section.

Between Classroom Analysis.

The zero order correlations between class mean science achievement and teacher behaviour indicate that the average class achievement in science is associated with a number of teacher and student behaviours. Classes of lesser ability tend to have teachers who contact more students about their work, use more rebuke of all types and more rewards of a non-verbal kind. Students in less able classes tend, on average, to spend less time listening, more time writing and are less likely to laugh or ask questions.

In general, relationships are not large and the most notable are the tendencies for teachers of less able classes to make more student contacts and make them write notes. The combination of these two behaviours, considered in conjunction with less time spent listening (without writing), appears to suggest that the activity occuring more frequently in less able classes is dictation or copying of notes. It is interesting to note that in addition to controlling the class by 'busying' students, more rebuke of all types is also more prevalent. The teachers also attempt to stimulate less able classes with non-verbal rewards, a system similar to that operating in many primary schools.

The degree to which the less able classes of students require these sorts of teaching techniques and the degree to which the techniques influence learning can be assessed by considering the beta coefficients. The lack of any statistically significant figures may be interpreted as evidence that the ability level of the class causes the teacher to respond in the way described above and not vice versa.

Between Classrooms Within School.

It is possible that the level of a particular teacher behaviour is, in part, a function of the influence or 'tone of the school and all the classes in a given school may be, for example, exposed to more rebuke than those in anothr school of differing tone. The degree to which the school tone associated with a particular teacher or student behaviour, is also related to science achievement would determine the extent of the correlation between the behaviour and achievement. The classes nested within school analysis effectively controls for the influence of the school and allows consideration of the effects of teachers who, for example rebuke more, or less, than is average for the teachers in their school.

Consideration of the zero order correlations indicate that the less able classes in a school appear to be given more dictation of notes and rebuked more than is average for that school. In addition to prompting the less able classes to work by traditional 'authoritarian' methods, teachers also appear to use more praise. However, it appears that some procedures are not used by teachers to stimulate such classes; namely student centred investigation. Students are less frequently invited to inquire, carry out less investigations and talk or read more to the class if they are in classes of less than average ability for their school. The only other correlations of significance indicate that there is less laughter and few invitations to consider work habits in relatively less able classes.

A comparison of the zero order correlations at the between classroom, and the between classroom nested within school, levels of analysis for some behaviours clearly indicate that the tone of the school influences teaching and learning styles. This can be noted for the following behaviours: students investigate, students talk, teacher invites students to inquire, teacher reviews work, teacher asks students to consider work habits, teacher praises, and finally, class laughs with the teacher. In these cases there is nothing systematic about the school's influence on the teacher's behaviour in terms of its relationship with the school's level of science achievement. On the other hand the following three variables appear to be systematically related to the level of achievement in the school: teacher uses non-verbal rewards, teacher uses casual rebuke, and students' laughter at the expense of the teacher or another student.

When student growth in science achievement is considered, four beta coefficients are 0.1 or above. It appears that classes which are rebuked more, or in which students carry out more investigation, or in which the teacher provides more positive support than is common for their school tend to improve more than would be predicted by their prescores, when compared to other classes in their school. However these tendencies are slight with the greatest per cent of variance in growth explained being 1.6 per cent.

Between School Analysis.

To check these interpretations, the correlations at the between school level must be examined. These correlations generally support the interpretations concerning the relationships between school tone and teacher behaviour described in the preceding section.

The standardized regression coefficients for this level of analysis show many values in excess of 0.10 and several considerably higher. The dangers of generalizing from only 15 schools are immediately apparent. Nevertheless a much publicized study recently released in Britain (Rutter, et al., 1979) has supported sweeping generalizations from only 12 schools.

SUMMARY

This study is of interest for a number of reasons. The first relates to the measures used to assess teacher behaviour and the comprehensive procedures followed to establish reliability. The second is associated with the large sample size which represents nearly the entire population. The third is concerned with the overall lack of relationship between teacher behaviours and growth in science achievement that is apparent from this secondary analysis. Although the first two of these features are relatively uncommon to research studies in the field of teacher effectiveness using observers, the last is far from unique. The lack of relationships between observed low inference teacher behaviour and student gain in achievement on cognitive outcomes has been noted frequently in reviews of the literature (one of the latest being, Rosenshine, 1978). It is, of course, possible to discuss the overall lack of relationship between teacher and student behaviour and science achievement gain as a consequence of unreliabilities of measurement of the independent and dependent variables. It is possible to take a more optimistic view and observe that the data reported here were collected in Canberra in 1969. The prevalence of streaming would have resulted in certain teaching behaviour being more highly correlated with both pre-test and post-test scores than would be found if streaming had not been present. In addition, the homogeneity of teaching behaviour imposed by an inspectorial system would have reduced the variance in the independent variables. Hopefully, a replication today would be more likely to identify effective teaching behaviour. However, one thing appears evident. Any investigation that does not exceed the one reported above in establishing the reliability and stability of its measures, or in the extent and scope of the study, is not likely to produce findings of greater strength than have been presented in this paper.

REFERENCES

ANTHONY, B.C.H. The identification and measurement of classroom environment process variables related to academic achievement. Unpublished Doctoral Dissertation, University of Chicago, 1967.

ARMSTRONG, J.R., VERE DE VAULT, M. & LARSON, E. Consistency of teacher communication: A sampling problem in interaction analysis. Psychology in the schools, 1967, 4(1), 74-76.

KEEVES, J.P. The home, the school and educational achievement. Unpublished Doctoral Dissertation, Australian National University, 1971.

KEEVES, J.P. The effects of the conditions of learning in the schools on educational achievement. IEA (Australia) Report 1974: 2. Hawthorn, Victoria: ACER, 1974.

MEDLEY, D.M. & MITZEL, H.E. Measuring classroom behaviour by systematic observations. In N.L. Gage (ed.). Handbook of research on teaching. Chicago: Rand McNally, 1963.

ROSENSHINE, B.V. Academic engaged time, content covered, and direct instruction. ERIC. ED 152776, 1978.

ROWLEY, G.L. The relationship of reliability in classroom research to the amount of observation. An extension of the Spearman Brown formula. Journal of Educational Measurement, 1978, 15(3), 165-180.

RUTTER, M., MAUGHAN, B., MORTIMORE, P., & OUSTON, J. Fifteen thousand hours. London: Open Books, 1979.

Research in Science Education,1979,9,23-31

A NEW STATISTICAL PROCEDURE FOR THE ANALYSIS OF HIERARCHY VALIDATION DATA

Alan R. Barton

INTRODUCTION

Early learning hierarchy studies relied on the use of various intuitive indices for deciding whether or not a postulated hierarchical sequence of skills was valid. The indices used included "proportion positive transfer" proposed by Gagné and Paradise (1961), a coefficient of reproducibility devised by Guttman (1944) and the phi-correlation coefficient. All of these indices were criticised by White (1974) because they failed to take account of errors of measurement and therefore lacked a sampling distribution.

White and Clark (1973) proposed a more sophisticated probabilistic model as a basis for testing the validity of postulated hierarchical connections. This model has attracted a great deal of attention from educational researchers and several major validation studies have been based on it (e.g. White, 1971; Linke, 1974 and Reeson 1977). However, White and Clark (and all other researchers who have used their model) have adopted a simplistic approach both to the estimation of the parameters of their model and to the actual test of the null hypothesis. The weaknesses of their approach are outlined below and a more effective procedure for the analysis of learning hierarchy validation data is proposed.

THE WHITE AND CLARK MODEL

Consider the following simple linear hierarchy:

The model assumes the existence of four mutually exclusive groups in the population:

P_0 is the proportion of the population who possess neither skill;

P_I is the proportion of the population who possess skill I only;

P_{II} is the proportion of the population who possess skill II only;

P_B is the proportion of the population who possess both skills.

To take into account errors of measurement the model also includes four additional parameters:

θ_a is the probability that someone with skill I will answer a question on skill I correctly;

θ_b is the probability that someone without skill I will answer a question on skill I correctly;

θ_c and θ_d are the corresponding skill II probabilities.

The hypothesis that all students who possess skill II also possess skill I is expressed as:

$$H_o: P_{II} = 0$$

The alternative hypothesis is:

$$H_a: P_{II} > 0$$

Using two questions to test each skill, the probability that any randomly selected member of the population will answer zero, one or two questions correctly for either skill is:

$$P_{xy} = \binom{2}{x}\binom{2}{y}[P_0\theta_b^x(1-\theta_b)^{2-x}\theta_d^y(1-\theta_d)^{2-y}$$
$$+ P_I\theta_a^x(1-\theta_a)^{2-x}\theta_d^y(1-\theta_d)^{2-y}$$
$$+ P_{II}\theta_b^x(1-\theta_b)^{2-x}\theta_c^y(1-\theta_c)^{2-y}$$
$$+ P_B\theta_a^x(1-\theta_a)^{2-x}\theta_c^y(1-\theta_c)^{2-y}]$$

where x and y are the numbers of questions correct for skill I and skill II respectively, that is, x = 0, 1, 2 and y = 0, 1, 2.

To obtain estimates for the parameters of the model the nine observed frequencies for each pair of hierarchically related skills were arranged in a 3 x 3 contingency table, with the number of questions correct for each individual skill represented by the appropriate marginal totals. The parameters were then estimated by setting up a series of equations which equated observed marginal totals with expected marginal totals.

Because the parameters were not uniquely determined from these equations for known marginal totals, White and Clark were forced to place restrictions on some of the parameters in order to obtain estimates for the others. The restrictions chosen were to set $\theta_b = 0$ and $\theta_c = 1$, with the claim that this was the most conservative choice since it minimized the probability of a Type I error. However, the values chosen for θ_b and θ_c are not estimates for these parameters but rather, setting $\theta_b = 0$ and $\theta_c =$

1 has the effect of removing these parameters from the model completely. Since these parameters were introduced in an attempt to account for errors of measurement, the claim by White and Clark that their test accounts for these errors is only partly true.

A further approximation was introduced by White and Clark to simplify computation. It was asserted that if H_a were true, the nine expected cell frequencies would be changed to some extent from their values under H_o, but the expected frequency in the critical cell (neither question correct for the lower skill and both questions correct for the higher skill) would have the greatest proportional change. Thus the test of the hypothesis $P_{II} = 0$ was based on the observed frequency f_{02} in this cell. The hypothesis was rejected whenever f_{02} exceeded a critical value, C. The critical value was determined by finding the value of k for which the cumulative probability

$$\Pr(f_{02} = 0) + \Pr(f_{02} = 1) + \ldots\ldots\ldots\ldots\ldots + \Pr(f_{02} = k)$$

first exceeded the critical value $(1 - \alpha)$, where α was the chosen significance level for the test.

Clearly there are three major sources of error introduced in this statistical procedure:

(i) it ignores the information by using observed marginal totals rather than the observed frequencies in individual cells;

(ii) it sets $\theta_b = 0$ and $\theta_c = 1$ in order to solve the complex equations equating observed marginal totals and expected marginal totals; and

(iii) it ultimately bases the test of the hypothesis $P_{II} = 0$ on the observed frequency in one cell (the so called critical cell) and rejects the hypothesis if the observed frequency in this cell exceeds a computed critical value.

MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS OF THE WHITE AND CLARK MODEL

Maximum likelihood estimation consists of selecting those values of the parameters θ_a, θ_b, θ_c, θ_d, P_0, P_I, P_{II}, P_B, for which the probability of obtaining the observed result is a maximum.

For any randomly selected member of the population there are nine possible outcomes associated with each connection in the hierarchy. Let the number observed in each of the nine classes be f_i $(i = 1,\ldots\ldots\ldots 9)$, and suppose that the probability that a single observation falls in the i th class is p_i $(i = 1,\ldots\ldots\ldots 9)$. Given a total sample of size n, the probability of obtaining f_1 in the 1st class, f_2 in the 2nd class and in general f_i in the i th class, is given by the multinomial probability distribution function

$$\frac{n!}{f_1!\, f_2! \ldots f_9!} \cdot p_1^{f_1} p_2^{f_2} \ldots p_9^{f_9}$$

where

$$n = \sum_{i=1}^{9} f_i$$

The likelihood function of the sample $\underset{\sim}{f} = (f_1, f_2, \ldots\ldots, f_9)$ is

$$L(\underset{\sim}{f}/\underset{\sim}{\theta}) = k[p_1(\underset{\sim}{\theta})]^{f_1}[p_2(\underset{\sim}{\theta})]^{f_2} \ldots [p_9(\underset{\sim}{\theta})]^{f_9}$$

where

$$\underset{\sim}{\theta} = (\theta_a, \theta_b, \theta_c, \theta_d, P_0, P_I, P_{II}, P_B)$$

and

$$k = \frac{n!}{f_1! \ldots f_9!}$$

The problem of maximum likelihood estimation is to select $\underset{\sim}{\theta}$ such that $L(\underset{\sim}{f}/\underset{\sim}{\theta})$ is a maximum. The constant k can obviously be ignored and so we are left with the problem of maximizing

$$L = L(\underset{\sim}{f}/\underset{\sim}{\theta}) = \prod_{i=1}^{9} p_i(\underset{\sim}{\theta})^{f_i}$$

An equivalent but simpler problem is to maximize

$$\log L = \sum_{i=1}^{9} f_i \log p_i(\underset{\sim}{\theta})$$

This can be achieved by solving the following set of simultaneous equations

$$\frac{\partial \log L}{\partial \theta_a} = \frac{\partial \log L}{\partial \theta_b} = \frac{\partial \log L}{\partial \theta_c} = \frac{\partial \log L}{\partial \theta_d} = \frac{\partial \log L}{\partial P_0} = \frac{\partial \log L}{\partial P_I} = \frac{\partial \log L}{\partial P_{II}} = 0$$

where

$$\frac{\partial \log L}{\partial \theta_a} = \sum_{i=1}^{9} \frac{f_i}{p_i} \frac{\partial p_i(\underset{\sim}{\theta})}{\partial \theta_a}, \quad \frac{\partial \log L}{\partial \theta_b} = \sum_{i=1}^{9} \frac{f_i}{p_i} \frac{\partial p_i(\underset{\sim}{\theta})}{\partial \theta_b} \quad \text{etc.}$$

and P_B is computed from the relationship $P_0 + P_I + P_{II} + P_B = 1$.

These partial differential equations are very complex. For example,

$$\frac{\partial p_1(\underset{\sim}{\theta})}{\partial \theta_a} = 2\theta_a P_I (1-\theta_d)^2 + 2\theta_a P_B (1-\theta_c)^2$$

$$\frac{\partial p_1(\underset{\sim}{\theta})}{\partial \theta_b} = 2\theta_b P_o (1-\theta_d)^2 + 2\theta_b P_{II} (1-\theta_c)^2$$

and so on, where $p_1(\underset{\sim}{\theta})$ is the probability that a randomlyselected member of the population answers both skill I questions correctly and both skill II questions incorrectly.

Because of the complexity of the equations, solution by analytic means is not feasible and thus iterative solution procedures must be used. A commonly used iterative technique for the solution of maximum likelihood equations is Fisher's method of scoring (Rao, 1965). In common with other iterative procedures the method assumes a trial solution and then derives equations for small additive corrections. The process can be repeated until the corrections become negligible.

When the method of scoring was applied to data reported in previous validation studies two serious problems were encountered. Firstly, the procedure failed to converge in some cases regardless of the intitial starting estimates. Secondly, in most cases at least one of the parameter estimates fell outside the acceptable 0,1 range.

These problems were overcome by using a "direct search" iterative procedure and a variable transformation which forced all parameter estimates to stay within the acceptable range. The direct search method which was used was the International Mathematical and Statistical Library routine ZXPOWL which is based on Zangwill's modification of Powell's conjugate directions algorithm. The variable transformation which was applied was to set

$$\beta_i = \sin^2 x_i \qquad i = 1,\ldots\ldots,7$$

where the β_i were the original parameters. Since

$$x_i \quad \varepsilon \quad \{\text{real numbers}\}$$

it follows that

$$-1 \leq \sin x_i \leq 1$$

and consequently

$$0 \leq \sin^2 x_i \leq 1$$

Once the $\hat{x}_i$ were obtained the $\hat{\beta}_i$ were computed from the relationship

$$\hat{\beta}_i = \sin^2 \hat{x}_i$$

Thus the $\hat{\beta}_i$ were restricted to the range

$$0 \leq \hat{\beta}_i \leq 1$$

as required.

THE LIKELIHOOD RATIO TEST OF HYPOTHESES INVOLVING P_{II}

Suppose we wish to test the hypothesis Ho: $P_{II} = 0$ against the alternative hypothesis Ha: $P_{II} > 0$.

The likelihood ratio test requires calculation of maximum likelihood estimates for the parameters θ_a, θ_b, θ_c, θ_d, P_O, P_I, P_{II}, P_B. These estimates are used to compute the unconditional maximum of the likelihood function

$$L = L(f_1, f_2, \ldots\ldots, f_9/\hat{\theta}_a, \hat{\theta}_b, \hat{\theta}_c, \hat{\theta}_d, \hat{P}_O, \hat{P}_I, \hat{P}_{II}, \hat{P}_B)$$

Then under the assumption that H_o is true, that is given $P_{II}=0$, maximum likelihood estimates for the remaining parameters are calculated, and used to compute the conditional maximum of the likelihood function

$$L' = L(f_1, f_2, \ldots\ldots, f_9/\hat{\theta}_a, \hat{\theta}_b, \hat{\theta}_c, \hat{\theta}_d, \hat{P}_O, \hat{P}_I, P_{II} = 0, \hat{P}_B)$$

The likelihood ratio statistic λ is then given by

$$\lambda = \frac{L'}{L}$$

λ is therefore the ratio of the conditional maximum of the likelihood function to its unconditional maximum or, alternatively, it is the maximum likelihood under H_o as a fraction of its largest possible value. The exact distribution of λ is not known but as $n \to \infty$, $-2 \log \lambda$ has a chi-square distribution with 1 degree of freedom (Kendall and Stuart, 1973, Vol.2, p.234).

When the likelihood ratio test was applied to hierarchy validation data reported by White (1971), Linke (1973, 1974, 1975) and Beeson (1977), using the levels of significance employed by the original researchers, the decisions reached differed from those obtained by these researchers for 30 of the 161 connections tested.

It must be noted, however, that the accuracy of the approximation of $-2 \log \lambda$ to the chi-square distribution has been theoretically established only when n is "large," although some confusion surrounds the meaning of

TABLE 1

Results of the Test of the Null Hypotheses for Validation Data which meets the Minimum Expected Cell Frequency Requirements

Scource of Data	Lower Skill	Higher Skill	Decision Using the White and Clark Test*	Decision Using the likelihood Ratio Test*		- 2 log λ
Linke	1/1(B)-V/S	2/1(B)-H/P	0.01	R	**	10.86
(1973; Vic)	1/2-V	3/2(A)-Max	R	R		36.04
	1/2-V	3/2(A)-Min	R	R		28.21
	1/1(A)-V/C	3/2(A)-Min	0.00	0.01	**	9.74
	4/1-V	4/1-V(Retest)	0.00	0.00		0.15
Linke	1/1(B)-V/S	2/1(A)-H/P	R	R		20.49
(1973; S.A.)	1/1(A)-H/S	2/1(B)-H/S	R	R		26.26
	1/1(B)-V/S	2/1(B)-H/P	0.00	0.01	**	9.41
	1/1(A)-V/C	3/2(A)-Min	0.00	0.00		3.95
	3/2(A)-Min	3/1	0.00	0.00		7.65
	4/2-V	4/1-V	0.00	0.00		1.06
Linke	1/1A	2/1B	0.00	0.02	**	9.45
(1974)	1/1A	3/2A	R	R		29.67
	3/2B	3/1	0.00	0.00		2.44
Linke	1/1A	2/1A	R	R		27.13
(1975)	1/1A	2/1B	0.00	0.05	**	8.94
	1/1A	3/2A	0.00	0.00		1.06
	3/2B	3/1	0.00	0.00		9.01
Beeson	XI	XXVa	0.00	0.01	**	9.89
(1977)	XI	XXVb	R	R		11.17
	XVIIIa	XIIa	0.05	0.05		5.29
	XVIIIa	XIIb	0.05	0.05		5.65
	XVIIIb	XVIIb	0.01	0.05	**	7.21

*Linke and Beeson each tested the null hypothesis at three different levels - the absolute level of no exceptions to the postulated connection and two weaker levels which allowed small but arbitary proportions of unqualified exceptions. The number represents acceptance of the appropriate null hypothesis. R represents rejection of all three null hypotheses.

** The likelihood ratio test and the White and Clark test lead to different conclusions.

large in this context and the accuracy may only be crucial when the computed value of $-2 \log \lambda$ is close to the critical chi-square value. The arbitrary rule which is normally applied is that expected cell frequencies should be at least five, although Conahan (1970) argued that the approximation is adequate provided the number of cells is at least five and expected frequencies in each cell are at least three.

In practice these recommendations are often ignored but, nevertheless, it seems sensible to view with caution results based on the likelihood ratio test when any expected cell frequencies are less than three. However, the data for 23 of the 161 connections tested met this requirement and the likelihood ratio test led to different conclusions from the White and Clark procedure for 7 of these connections. The results are listed in Table 1. Thus, even if the conservative position is taken, conclusions based on the likelihood ratio test differ markedly from those based on the White and Clark test. The proportion of differences (30 out of 161 or 7 out of 23) are both well outside the stipulated α level (probability of a Type 1 error) used in the studies cited.

CONCLUSION

The likelihood ratio test is a powerful and soundly based statistical procedure which can be applied to the White and Clark model without introducing *ad hoc* approximations or assumptions. Since use of this test often leads to different decisions than those reached using the procedure suggested by White and Clark, it is recommended that it be used in future hierarchy validation studies.

REFERENCES

BARTON, A.R. A simple technique for improving the quality of parameter estimates in learning hierarchy validation studies. *Psychometrika*, 1980, in press.

BEESON, C.W. Hierarchical learning in electrical science. *Journal of Research in Science Teaching*, 1977, 14(2), 117-127.

CONAHAN, M.A. The comparative accuracy of the likelihood ratio and X^2 as approximations to the exact multinomial test. Unpublished doctoral dissertation, Lehigh University, 1970.

GAGNÉ, R.M. & PARADISE, N.E. Abilities and learning sets in knowledge acquisition. *Psychological Monographs*, 1961, 75(14), 1-23.

GUTTMANN, L. A basis for scaling qualitative data. American Sociological Review, 1944, 8, 139-150.

INTERNATIONAL MATHEMATICAL AND STATISTICAL LIBRARY. Library 3. 4th ed. Houston, Texas: IMSL, 1974.

KENDALL, M.G. & STUART, A. The advanced theory of statistics, Vol. 2, 3rd ed. London: Griffin, 1973.

LINKE, R.D. The effects of certain personal and situational variables on the acquisition sequence of graphical interpretation skills. Unpublished doctoral dissertation, Monash University, Victoria, 1973.

LINKE, R.D. Influence of cultural background on hierarchical learning. Journal of Educational Psychology, 1974, 66(6), 911-918.

LINKE, R.D. Replicative studies in hierarchical learning of graphical interpretation skills. British Journal of Educational Psychology, 1975, 45, 39-46.

RAO, C.R. Linear statistic inference and its applications. New York: Wiley, 1965.

WHITE, R.T. Learning graphical skills in kinematics. Unpublished doctoral dissertation, Monash University, Victoria, 1971.

WHITE, R.T. Indexes used in testing the validity of learning hierarchies. Journal of Research in Science Teaching. 1974, 11(1), 61-66.

WHITE, R.T. & CLARK, R.M. A test of inclusion which allows for errors of measurement. Psychometrika, 1973, 38(1), 77-86.

Research in Science Education,1979,9,33-41

SCHOOL-BASED DEVELOPMENTAL RESEARCH—FROM IDEAS TO IMPLEMENTATION

J. R. Northfield and L. C. Ingvarson

THE DEVELOPMENT OF IDEAS ASSOCIATED WITH SCHOOL-BASED DEVELOPMENT

The literature on teacher education and curriculum implementation has been marked by an increasing use of terms which reflect greater teacher autonomy in three areas of education. Phrases such as "school-based curriculum development," "teacher as evaluator," "teacher as researcher," "action research" reflect a recent trend and are now part of the jargon associated with teacher education and curriculum development. Schools and teachers have become more responsible for initiating and implementing their own in-service and curriculum development activities and a number of reasons can be proposed to account for this shift in responsibility.

A first group of reasons for greater teacher autonomy in important areas of education can be described as a response to dissatisfaction with the limited impact of curriculum developments and in-service opportunities which have existed. School-based in-service activities may provide alternative approaches likely to be more relevant to the classroom. This view that the education "industry" outside the schools is not affecting schools as much as might be desirable is a very important factor in the arguments for school-based approaches. This group of reasons will not be explored further in this paper, although recognized as indicating an important, complex and debatable issue.

Secondly teacher autonomy is consistent with a wider acceptance of diversity among schools and a need to develop a closer link between schools and the communities within which they operate. Curriculum and teacher education needs are likely to vary widely and the schools via their teachers are ideally the best people to initiate and implement the most appropriate activities. Shifting the responsibilities to the teacher means that the teacher is also being accepted more as a professional person capable of initiating and carrying out tasks related to his/her own teaching and development. These responsibilities and "new" roles have not been easy for teachers to adopt, a point which will be developed further later in the paper.

Thirdly, recent research on the problems of implementation (Fullan & Pomfret, 1977) have influenced recent thinking about in-service education. Local variability in implementation of a curriculum programme is viewed as inevitable and a sign of significant and desirable change to suit the local setting. In-service is therefore required during the implementation process rather than a programme of activities prior to implementation. The major focus is on development of the teacher rather than dissemination of information about a curriculum approach. To quote one view:

> ..a most important aspect of significant change is not so much the "wheel" or the educational technology but the process of "reinvention" or individual development(McLaughlin, 1976 p.351).

The factor associated with the trend to greater teacher autonomy that will receive greater attention in this paper is the increased support for in-service education as a result of the Australian Schools Commission policies. Two major themes concerning in-service education pervade the reports that the Australian Schools Commission has produced since its inception in 1973. These quotations are representative:

> The Commission reaffirms that, through the (INSET) Program, the teaching profession should take a leading part in directing its own improved functioning... (Schools Commission 1978, p.89).

> Perhaps the most important single grouping for developmental activities is the total group involved in the work of the school (Schools Commission 1975, p.185).

The reports have advocated that in-service education should largely be an area for teacher initiative, responsibility and decision-making. In addition the reports suggest that the "Development Programs" which the Commission fund should aim to break down the professional isolation of teachers in particular schools by providing opportunities for school staff (as a group, or sub-group) to take in-service education initiatives.

The Commission has placed its confidence squarely in the school as the most significant unit for change and teachers as a most appropriate source of ideas for developmental activities. The means by which the ideas were to be translated into practice was for each state in-service education committee to determine.

A number of surveys have indicated strong teacher support for school-based teacher developmental activities (Batten, 1979), but despite this support it has proved difficult to develop suitable approaches to school-based development. In particular, the two features of being school-based and teacher-initiated have been rarely combined successfully.

THE DEVELOPMENT OF SCHOOL-BASED IN-SERVICE IN VICTORIA

After five years the Schools Commission's INSET programme appears to have affected schools and teachers in a number of ways. The INSET Program has involved a large number of teachers and has made the administration of in-service education within Victoria much more transparent to teachers in schools. What was formerly a mystery has become clearer. More teachers know how to get help in organizing a program and how to make a submission to a Regional or State in-service education committee for financial support. The administrative structure of the INSET Program has been designed to facilitate activities which reflect the expressed needs as perceived by teachers. Assessing needs has always been difficult and the many questionnaire surveys used have rarely provided useful guidelines. In a sense the Program has catered for this role through its smorgasboard or responsive attitude. This assumes that a Program based upon a wide range

of submissions from potential in-service organizers will represent the major areas of in-service interests of teachers.

The extent to which the Program does in practice reflect teacher needs is a contentious issue and there are indications that the Victorian state committee is shifting from a responsive role to a role of initiating different types of activities. The major emphasis of the program in the initial years has been on dissemination of ideas. Whether the activity was a course or a conference, in or out of school it has generally been a case of telling, advising, showing teachers 'how to do it' rather than finding out with them 'how it might be done.' It is rarely possible for in-service activities of this type to help teachers put ideas into practice. There was a need to develop activities which assisted teachers to implement ideas rather than stopping at the disseminating function. The national evaluation report expresses the impact of the Program on teachers in this way.

> It is four years since the Development Program commenced, and time for development committees to give serious consideration to the direction the Program should take in the future. There are two important factors which are germane to such considerations:
>
> (i) Teachers have appreciated the influx of in-service courses that have been offered to them, but they have now become a little blase and less tolerant of mediocrity. Committees and organizers can no longer be excused on the grounds of inexperience for approving or producing second-rate stereotyped programs.
>
> (ii) As demonstrated in the Victorian questionnaire survey, the most important impetus for change in teaching behaviour is self-motivation. In-service education would become a much more powerful force if it could look into this highly personalized aspect of professional development (Batten 1979, p.166).

A clear indication for the direction the Program might take in the future is provided later in the report.

> No single aspect of the Program has been received with such universal acclaim as the trend toward school centred professional development. The Commission, development committees, teachers, administrators, parents, consultants, teachers' centre personnel...all these people have expressed support for the idea and the reality of school centred development work. The reason may be that it incorporates so many of the factors which have been important contributors to effective in-service education within the Program. The school centred element represents the final move in the strongly supported regionalization trend; it is an element which favours active participant involvement and in-service continuity; it presents an opportunity to benefit an answer to the perpetual cry from teachers that in-service work should be directly relevant to their work in the classroom (Batten 1979, pp.287-8).

At the administrative level three states have set up innovations sub-committees "to play a different and more positive role in the Development Program in addition to their major accepted role as

'response committees'" (Ratten 1979, p.152).

This indicates a commitment to exploring new approaches rather than relying on proposals put forward for funding. By 1979 the school-based in-service activities in Victoria can be grouped in the following way:

(i) In-service activities which focus on school organization issues. e.g. school reviews and programmes involving the wider school community

(ii) Activities which have a more restricted focus on one area of the curriculum e.g. single subject reviews within the school

(iii) Activities designed to facilitate individual/small group development in certain aspects of teaching.

The remainder of the paper will describe an example from this category "The Classroom-based Developmental Research Project." This is a project suggested by the Program Initiatives sub-committee in Victoria. It is an attempt to incorporate the two ideas of school-based and teacher initiated, two themes often expressed but rarely combined in a practice.

TRANSLATING THE IDEAS INTO PRACTICE

The clearest expression of how the ideas associated with classroom-based professional development might be translated into a practical approach is to be found in the Ford Teaching Project (Elliott & Adelman 1976). This project appears to be a practical extension of the ideas of Stenhouse (1975) and the experience gained during the development of dissemination of the Humanities Curriculum Project (Rudduck 1976, Elliott & MacDonald 1975). These ideas provided the most promising guidelines for the INSET shift towards teacher development by making the study of teaching by teachers the focus for in-service education activities.

Although providing a clear model of "action research" as the term is defined by Elliott (1978) the Ford Teaching Project could not be regarded as teacher-initiated. The initiative in the British program was with the researchers not the teacher. In March 1978 VISEC provided funds to support a project which was to be teacher initiated as well as school based (Classroom-based Developmental Research Projects 1978).

Fifty submissions were received from groups of teachers from all types of schools. The call for applications from teachers described the project in the following way.

> The main purpose of this form of in-service education is to support classroom-based research studies by groups of teachers interested in exploring the effectiveness of teaching programs or methods they are currently using or planning to implement. It is felt that such case studies carried out by teachers under ordinary working conditions will be of interest and value to other teachers (Education Department, Victoria, 1978).

Eight proposals were selected for support throughout 1978 and the teacher groups met together for the first time in May. This meeting provided the first opportunity for each group to explain their study to other teachers. Each teacher group had invited an outside "consultant" to work with them. The teachers and consultants discussed the methods of gathering the information that might assist in the understanding of their particular topic areas.

Two further meetings were held with the teacher groups throughout the year. Apart from the expectation that each group would discuss the progress of their project, the meetings were unstructured. The time spent in meeting together proved to be inadequate. The groups became very interested in projects other than their own. Teachers were keen to canvass the opinions of other teachers. Problems were freely expressed and discussed and in each case the original plans and proposals had been extended to include consideration of much broader educational issues. The initial study had opened up issues and questions and the teachers were committed to further classroom study. By the third meeting four of the groups expressed the intention of requesting an extension to continue with their work in 1979. In each of these cases little in the way of extra funds was required. The main purpose of the extension was expressed as "an opportunity to maintain contact with a group of teachers who are interested in hearing about the project and willing to react to our ideas."

The final section introduces some of the issues and problems that arise in action research of this type. The examples will be drawn mainly from the experience with the project outlined above although the issues and problems discussed appear to form a consistent feature of other action research efforts in the authors' experience.

THE PROBLEMS ASSOCIATED WITH "SCHOOL BASED TEACHER DEVELOPMENT"

Reporting the Results of Each Activity.

In November 1978 each group of teachers prepared interim reports for VISEC. In many respects these reports were a little disappointing. The reports failed to provide convincing evidence of "professional development" for a committee now asked to consider applications for further support for four projects as well as whether to call for further applications for support in 1979. What little thought given to evaluating the overall project had assumed that the group reports would be a major piece of relevant documentation. There appear to be several reasons why the participant reports did not reflect the value of the project and many of the issues and problems of this form of research seem to be closely associated with the way in which the activities are reported.

The reports appear to have two main purposes, an evaluation function and a less well defined communication function to a wider audience of teachers and educators. The evaluation purpose allows the sponsors to monitor the project and assess the priority which should be given to Classroom-Based Developmental Research in comparison with other VISEC activities. With limited resources and a growing number of applications for in-service support this evaluation function must be maintained but an

over-reliance on the written reports of participants may be an inappropriate way of monitoring the overall project. The participants expressed concern about the form of reporting required from the start of the project. In all but one case the teacher reports did not fit the description described by Elliott (1978).

> Since action-research looks at a situation from the participants' point of view it will describe and explain "what is going on" in the same language as they used; namely the commonsense language people use to describe and explain human actions and social situations in everyday life (Elliott 1978, p.356).

In general the teachers have attempted to write for the sponsor of the project rather than in the way they described their studies to other teachers at the meetings throughout the year. In future it will be necessary to give teachers more guidelines for reporting in a more informal case study style. The sponsors will need to consider other ways of evaluating the overall project. Perhaps evaluating the degree of teacher development that is occurring requires discussion with the participants and their colleagues and pupils in the schools.

The second function of the reports is clearly illustrated by reference to the notice calling for applications.

> A report will be required by VISEC at the completion of each development research project so that these practical and detailed case studies might be made available to other interested persons or groups. (Education Department, Victoria, 1978).

There appears to be an assumption that the case study might provide "a surrogate experience" which "invites the reader to underwrite the account by appealing to his tacit knowledge of human situations" (Adelman et al. 1976, p.143). Fensham and Ingvarson (1977) have suggested that this use of the case study may lead to reader to generalize from one situation to his/her own more familiar situation:

> If information is presented in a way that enables other persons (who are quite outside the situation from which the information is drawn) to see relevance for their own behaviours, then this information has generalizability (Fensham & Ingvarson 1977).

Such a concept means that any generalizations are to be large extent "generated by an audience other than that which initiated the research (Fensham & Ingvarson 1977).

This communication function of the reports is an important part of the activity. Although only a small number of teachers can be supported directly there is an assumption that the reports of their activity might have a wider impact on teachers. Unfortunately we seem to know very little about the types of reports that will have an impact on teachers let alone whether the personal accounts of teacher experience are likely to be read by other teachers.

To a large extent the preliminary evaluation of the project is related to the issue of reporting each of the funded activities. The teachers have engaged in a very personal process as part of their normal teaching. It is

difficult for them to find the time to record the progress of their classroom research. From discussions with teachers there is evidence of confusion about the purposes of the report. They feel unable to approach the task of reporting to the administrators and express doubts that other teachers would want to read of their unique approaches to often unique problems. There is no doubt that this group of teachers needed greater assistance with the preparation of their reports and reassurance that their contributions would be of value for other teachers.

The Importance of Interaction with Colleagues.

The experience of one year with the VISEC project has emphasized the importance of interacting with other teachers, a factor clearly evident in earlier studies of in-service needs.

> The factor that emerged as by far the most important for professional development was contact with other teachers: simply exchanging ideas; observing other teachers' lessons; working with experienced staff; visiting other schools... (VISEC 1976, p.14).

As already indicated three meetings were little more than a token effort in this regard and perhaps the professional development of the participants was limited by the failure to organize opportunities for greater teacher-teacher interaction.

A requirement to discuss their research with other teachers can have a number of effects on a group of teachers. It can provide a deadline to collect together ideas and discuss their progress...a reason for putting aside the day-to-day teaching activities and reviewing the existing situation. The need to clarify and express ideas to other teachers can be regarded as an essential part of the learning experience for teachers. The opportunity to hear reactions and suggestions from others might provide new ideas and perspectives on "the research task." Teachers seemed to react very favourably to the sessions when the groups met together, despite the fact that they were held in the late afternoon and the time spent on the work of any one group was very limited. The requirement for greater interaction between groups should be made a condition of involvement in further programmes of this type. More time should be available for each group to receive support, ideas and advice from colleagues and consultants. The time to reflect, review and plan ahead may not be available in the school yet it appears to be crucial for the progress of classroom research of this type. The request of four groups that they continue to meet in 1979 certainly emphasizes the importance of teacher-teacher interaction.

The Time Required for Teacher Development.

The time limitations for a teacher engaged in action research become apparent in two ways. Two quotes from another action research project seem to convey the effects of limited time. In the first case we have the difficulty in finding time to engage in the research, write reports, etc.

> I think I have begun lots of useful things but I certainly have not got it all together or followed it through. One of the major reasons

> for this I think is because so many other things have been happening at the school. I think I could have achieved more if life had not been so hectic generally...I frequently found myself resenting the extra pressure which this project inflicted on me...

In the second quote the teacher seems to be referring to the time required to develop an understanding of the area of study.

> Another problem was the frustration of the limited time we had together... there was so much to get through and so, as a consequence, too much was left unresolved. I wanted so much to share with others my experiences and learn from them, but, instead of a relaxed joyful learning experience, I often took home with me a headache... how disappointing; a creative experience needs time.

Despite the evidence available from a growing number of action research projects we still seem to think of providing support for too short a period of time. At the end of the first year (about 8 months of teacher activity) most groups were only beginning to identify and understand their original problem area. This problem area had also grown to include related issues not considered in the initial planning. The activities of one group can perhaps serve as an example. This group of teachers received support for a project designed to develop oracy skills across all subjects of the curriculum. The original intention was to experiment with ways of developing and assessing the speaking and listening skills of pupils. To analyze the meaning of the term "oracy," develop or justify approaches to developing speaking and listening skills in different subjects and develop appropriate methods of evaluating these skills, is a very complex task. It was inevitable that this group of teachers would have to consider issues of curriculum evaluation (justifying and describing what they were doing to administrators, project sponsors etc.), approaches to in-service education (involving other staff in the project) and communication with parents of their pupils. The initial support for the project had only allowed sufficient time for the problem area to "unfold" for the participants. A similar pattern occurred for each of the other projects. In future it will be necessary to either find a way of restricting the scope of action research studies or allow for support to occur over longer periods of time as different problems and issues emerge. It would seem a pity if the former alternative is followed. The problems and issues which grow out of studies such as this tend to be important issues for teachers which the action research study seems to bring to their attention. The study seems to mean that teachers are forced to confront some of these education issues and perhaps more clearly understand and solve some of the problems.

The classroom-based developmental research project can be regarded as an experiment in in-service education with exciting possibilities. As with other models of action research there are a number of problems emerging when the ideas are implemented. Just as we ask teachers to engage in a process which may lead to a deeper understanding of their teaching, the organizers of action research programmes must engage in a monitoring of their programmes. The practical difficulties can be too easily overlooked as the trend to school based curriculum development and teacher education becomes more evident.

REFERENCES

ADELMAN , C., JENKINS, D. & KEMMIS, S. Rethinking case study: notes from the second Cambridge Conference. Cambridge Journal of Education, 1976,6, 139-150.

BATTEN, M. Report of a national evaluation of the development program. Schools Commission Evaluation Studies. Canberra: Schools Commission, 1979.

EDUCATION DEPARTMENT, VICTORIA. Education gazette and teachers aid. 1978, 128(2), 68.

ELLIOTT, J. What is action-research in schools? Journal of Curriculum Studies, 1978, 10(4), 355-357.

ELLIOTT, J. & ADELMAN, C. Innovation at the classroom level. Unit 28 of the Open University Course Curriculum Design and Development. In Innovation, the school and the teacher (1). Open University Press, 1976.

ELLIOTT, J. & MacDONALD, B. (Eds). People in classrooms. Centre for Applied Research in Education. University of East Anglia, 1975.

FENSHAM, P.J. & INGVARSON, L.C. Case study research in science classrooms. Paper presented to the Conference of the Australian Association for Research in Education, Canberra, 1977.

FULLAN, M. & POMFRET, A. Research on curriculum and instruction implementation. Review of Educational Research, 1977, 47(2), 335-397.

McLAUGHLIN, M.W. Implementation as mutual adaptation: change in classroom organization. Teachers College Record, 1976, 3, 339-351.

RUDDUCK, J. Dissemination of innovation: the Humanities Curriculum Project. Schools Council Working Paper 56. London: Evans/Methuen Educational, 1976.

SCHOOLS COMMISSION. Report for the triennium 1976-8. Canberra: Australian Government Publishing Service, 1975.

SCHOOLS COMMISSION. Report for the triennium 1979-81. Canberra: Australian Government Publishing Service, 1978.

STENHOUSE, L. An introduction to curriculum research and development. London: Heinemann, 1975.

VISEC. School-based teacher development. Victorian In-Service Education Evaluation Report. Glenbervie Teachers Centre, 1976.

Research in Science Education,1979,9,43-52

THE FOREST AND THE WOODCHIPS—ARE THEY ALTERNATIVE PARADIGMS FOR STUDYING THE TREE OF SCIENCE EDUCATION?

Peter J. Fensham

INTRODUCTION

Power (1976) alerted the science education communities in Australia and in the U.S.A. to the existence of what he described as "competing paradigms" for research in science education. These can be described as the traditional or "agricultural - botany" one and the other, the case study or anthropological one. Hamilton (1978) has contrasted these two approaches and distinguishes them from experimental studies in which the natural environment of schools and classrooms is not maintained, and from simple census surveys of direct factual data.

In Britain where the "agricultural" paradigm was less established and practised, the "case study" approach was already better known through the influential work of Hamilton (1975) in Scotland and others like Walker (1974) and Elliot (1973-75). These scholars had been active in evaluative studies of some of the Nuffield science curricula and the Ford Teaching Project with its concern for an inquiry approach, a pedagogy that was strongly intended in a number of the English science curriculum projects at both the primary and secondary levels.

Since Power's paper there have been several attempts (e.g. Suan, 1976; Fensham & Ingvarson, 1979) to use the second research approach in Australian science education, and the publication of the very large research project, Case Studies in Science Education, by Stake and Easley (1978) has now made this type of research much more visible to science educators. It is thus appropriate that critical scrutiny be made of these two approaches to research in science education since many people have been rather confused by the rather polemical nature of much of the writing about the relative merits of these two kinds of research. The use of the word "paradigms" to describe them has been seen by some to be inappropriate, and softer distinctions such as "alternative" or "complementary" have been suggested. The popularity of the ideas of Kuhn (1962) concerning normal and revolutionary science to describe the historical development of scientific knowledge has encouraged both a loose use of the term "paradigm" to describe alternative theories, models or techniques of research, and the radical suggestion that something much more fundamentally different exists between these two approaches to research into science education.

Good educational research is not easy to carry out within either of these approaches. It has been easy for critical proponents from either camp to point out serious and quite disabling weaknesses in the design and conceptualisation of research studies in their own or the other camp. Gardner (1974), for example, did traditional research in science education a great service by enumerating the sorts of basic errors that havea been perpetuated in many studies of attitudes in science education and by suggesting a number of ways to improve such studies. From a similar

position, Parsons (1976) pointed to a number of weaknesses in studies using the case study approach. Parlett and Hamilton (1972), from the other side, have been harsh critics of the triviality that is implicit in the starting points and the resultant studies of traditional researchers. One reviewer of the C.S.S.E. said "it would be for educational studies like Lewis and Clark's report - a hallmark for future work." Another said "enough paper had already been wasted on it."

SOFTER RESPONSES TO THE TWO APPROACHES

My own original introduction to case study research took place in a critical, but less potentially polarised situation. In the mid-50's there was a rash of research in social psychology using the traditional approach. This was not without its critics, and Ittelson and Cantrill (1954) and Zangwill (1950) were the guides for my own beginnings. The former, at Princeton, were reporting findings that stressed the subjective and interactive nature of social perception, and hence the importance of context in social research. The latter saw the use of all the components of the agricultural-botany approach (definition of variables, hypothesis, measurement, testing of hypothesis, etc.) as premature for this field. Social psychology (or paraphrasing it to our interest, science curriculum research) should be content, at this stage, with long periods of rather open ended observation and exploration of the phenomenon under study. Only after such study, as in the more developed sciences, would variables of real significance emerge. Thus, it was only at the end of my own 2-year long participant observation that I tentatively suggested the proposal (hypothesis) that the changing social cohesiveness of the groups in an innovatory situation was the key to outcomes of successful or unsuccessful adoption (Fensham and Hooper ,1959). This hypothesis between variables defined from long study of the situation could, in principle, then be tested in a survey analysis study of the traditional type across a number of similar situations. In fact, such a survey could not have been carried out on this hypothesis without a case study of every instance, because both my variables were only operationally definable in terms of each unique situation. Nevertheless, there did not seem to be a question, in Kuhn's sense, of there being two paradigms. The difference in the research approaches was in their sense of the stage of development that had been reached in the same "normal" view about how social phenomena can be explained.

Cronbach (1975) and Campbell (1975) in the U.S.A. have tried in recent writings to maintain the unity of the research community by subsuming the case study into traditional research as a study of a population or as a pre-experimental design (shades of my own earlier inclination above). However, the English groups of case study researchers show little real interest in such compromise.

There is an attractiveness about ideas like "we are not divided," that science education research is not in a crisis, that somehow we can find complementary or compatible alternatives in the two approaches. However, I wish to argue in this paper that Power was formally correct and not loose in his use of "paradigm" to describe the two approaches to research on natural science classrooms even though he himself hoped they were alternatives.

The discussion that follows does not relate to research in which the design sets up experimental classrooms. These seem quite appropriate for investigating many questions that are basic to science education, but their findings have other serious problems concerning how to relate them to natural classrooms.

THE CASE FOR TWO PARADIGMS

Kuhn in his very influential book of 1962 says "scientific revolutions are inaugurated by a growing sense, often restricted to a narrow subdivision of the scientific community, that an existing paradigm has ceased to function adequately in the exploration of an aspect of nature to which the paradigm itself had previously led the way." However, just what he meant by the word, paradigm, in this book was not easily discerned by many readers. Masterman (1970) for example, pointed out that the word is used by Kuhn in at least 22 different ways. In his Postscript to the second edition of his book, Kuhn (1970) sets out to define more clearly what the term does mean to him and it is the following, more distinct characteristics that I will use in the argument of this paper.

I. A paradigm governs in the first instance not a subject matter but rather a group of practitioners.

In Britain there is now undoubtedly a group of case study researchers who see themselves (and are seen by the others) as different from other educational researchers. In some universities most if not all the curriculum researchers belong to this group. Their strength as a group - apart from their own research publications - is evident in their domination of the Open Books series on curriculum and in the curriculum materials for the Open University. In the U.S.A., the visible acknowledgment of the group as distinct is probably less, but they know and trust each other, share mailing lists, and convene regularly in meetings to share examples; even though these may be less formal than those that have occurred in England at Cambridge. Although there are considerable differences within the case study group as is evident if one compares Stenhouse (1978) and MacDonald and Walker (1975), their one-ness, in comparison to traditional researchers, transcends these differences.

II. A paradigm is what the members of a scientific community share.

The traditional researchers have much they share. There is a well established set of methods of investigation, designs for research, concepts worth studying, and findings that stimulate replications and other studies extending the research that is reported. There is a commitment to these things that drives their research and governs their discussion of the findings. Phillips (1974), at an earlier meeting of this association described these aspects as a conservative influence on their research. Even among highly respected members of the community one finds a strange allegiance to beliefs about what their research design and analysis achieves.

Two examples will suffice. Nicodemus, Jenkins and Ingle (1976) reported on a survey analysis they carried out concerning the adoption of Nuffield chemistry by teachers and schools in England. Data was gathered on about 100 different teacher characteristics from 51%(!) of their teacher sample. Fifteen percent of these characteristics were found to be significantly different for high adopters and low adopters. The latter group contained rather strangely the teachers who said they "used the ideas of Nuffield but not the materials" - a curious categorisation in the face of the strong pleas by English writers that a curriculum is "principles ... for translation into practice" (Stenhouse, 1975) or "an acknowledged framework of norms and decisions" (Eraut, 1978). The paper reporting this study goes on to devote all its remaining paragraphs to these few differences and ignores the overwhelming evidence that these teachers are the same. Indeed so strong is the focus on statistically significant relationships between variables that, with respect to the cluster of characteristics about sources of information, two discriminating sources are discussed even though the teachers, as a group, had not rated either of them among the first four most frequently used sources.

Rothman, Walberg and Welch (1968), another highly respected group of traditional researchers, in like manner in a study of the effects of inservice courses for physics teachers, ignored completely significant differences in the sub-cluster scores making up a non-significant composite variable, and then proceeded to discuss the meaning of the overall significance of their other composite variable in terms of its significantly different sub-cluster ones.

In the other camp, Stake commissioned eleven case studies for C.S.S.E. and published eleven. Although it seems likely that the quality of these must have been rather variable, readers are not assisted by the project directors' parts of the report to identify which findings should be given greater weight than others. The findings of the brief site visitors also appear alongside extracts from the case records of the much longer "resident" researchers, and seem to carry considerable weight in the general findings.

The common sharing of the case study community is less easy to list but it is easy to experience if one spends time with them. They do share methodologies, although at first sight these seem often to be vague and disparate. They share findings although to the uninitiated what is shared is not obvious since they are not neatly summarised as in the traditional community. Good case studies are shared in the way Kuhn describes when he talks of paradigms as shared examples. Although they may describe quite different educational "problems" or situations, members of the case study community will read them as being about the same sorts of reality. Perhaps most obvious as a sharing when you move among them is the common belief about research and reality.

Case study workers are a group of researchers who, as Kuhn says, "have a common awareness that something is wrong" with traditional research in education. Much of their writing is concerned with this but, to outsiders, this can easily be confused with the constructive criticism of authors like Gardner (1974) or Snow (1974).

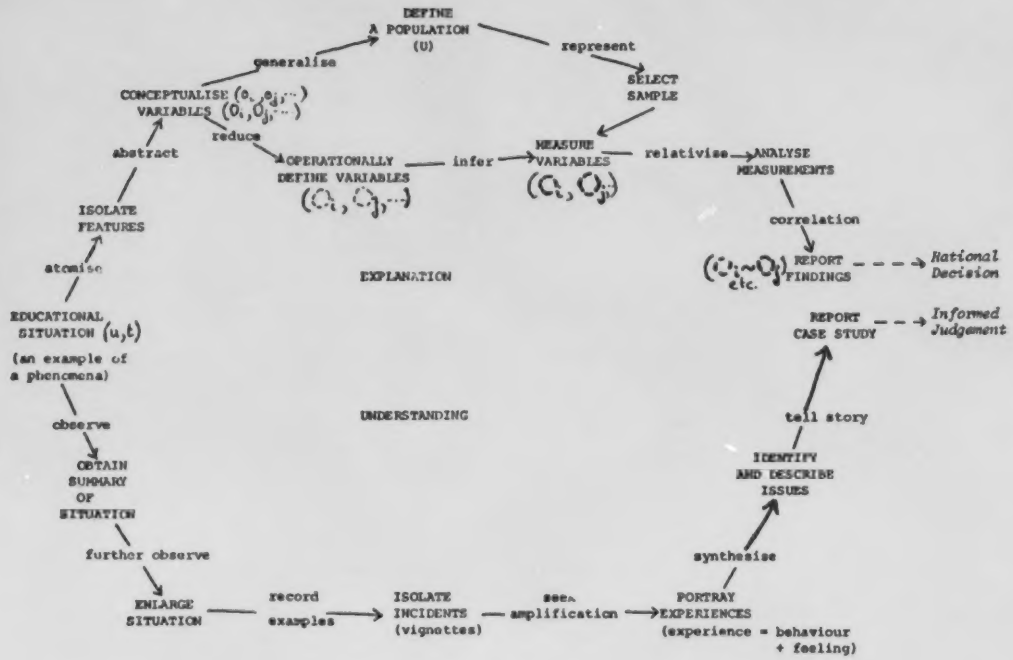


Figure 1: The steps in the two research approaches - upper path, traditional research; lower path, case study research

However, their concern, say with the main findings Rosier (1974) reported from the I.E.A. study of science education in Australia, would not be that they are trivial because our research techniques are not yet sophisticated enough for us to establish more penetrating generalisations. Rather, it is that such studies and their findings obscure the reality of what science classrooms are like in Australian schools.

III. Paradigm is a constellation of shared commitments.

The commitments of the two communities have been written about by Stake (1972) in his essay on Description versus Analysis, but became much clearer and more cogent to me in informal discussions with him during 1978. Figure 1 is an attempt to illustrate the nature of the different commitments.

Let us start with a real phenomenon in science education, e.g. a classroom where a science teacher is attempting to teach an intended curriculum or the diffusion to schools and classrooms of a new curriculum from a project or curriculum authority. The upper loop in Figure 1 is the pathway of traditional research.

The traditional researcher observes a situation (u) an instance (or instances) of the phenomenon (t) and embarks on a process of explaining it, using Cronbach's notation for instances u (population U), for observables o (O) and for treatments, t (T). To do this he isolates from the phenomenon a number (often very few) of features that to her/him seem likely to be significant (a process of atomisation). These are then conceptualised (o_i, o_j) and defined as variables (O_i, O_j)(a double process of abstraction) that can now apply in other similar situations, u_i (e.g. science classrooms) within a population, U. By now, the starting situation is no more than but one example, and not the unique situation its participants - teacher and learners - know it to be. A relationship between pairs or clusters of these variables ($O_i \sim O_j$) is postulated by the researchers to hold over the population. To test this hypothesis, a sample (Σu_i) is drawn from the population which now need not even include the situation from which this process started. The variables are now operationally defined for measurement purposes (O_i, O_j) and their values (O_i, O_j) collected across the sample (Σu_i). A variety of statistical procedures exist from which a method of analysis is chosen which best matches the assumptions and nature of the measures (parametric, non-parametric etc.). In this way some of the hypothesised relationships are found to hold in a statistical sense, say $O_i \sim O_j$, while others do not. The researcher then either <u>explains</u> the original phenomena in terms of the variables (O_i, O_j) as if they are casually related or suggests that the <u>explanation</u> of the situation will lie in further study of this significant relationship. A very well designed example of this type of research in Australian science education is the recent study of the diffusion and adoption of ASEP. The authors (Owen & Tisher, 1978) draw the more cautious conclusion above, and suggest that case studies of exemplary schools, from the clusters in which different significant relationships hold, are the next steps for research.

With at least three processes of abstraction and two steps to lose uniqueness, this traditional approach is committed to a belief that the social behaviour of educational situations is an outward manifestation of

underlying, more fundamental properties, that transcend the confines of particular participants and their unique contexts. This belief about the atomistic basis for social behaviour is akin to the chemist's view when he accepted ionised atoms and molecules as the common basis of the varieties of conducting solutions that can be formed from all sorts of solutes and solvents.

The pathway of the case study researcher is indicated by the lower part of Figure 1. He/she looks at the situation and decides that it requires a lot more observation because it is complex and because only some impressions are obvious initially. Longer observation (or conversation if, like Stenhouse (1978), you favour interviewing over participant observation) lead to incidents ("vignettes" in Stake's terms) about which clarification or amplification is sought from the participants. The experience the participants have of the incident will often reveal that it is linked to wider aspects of this situation. Repeated amplification over time of such incidents reveals more clearly the nature of these more total aspects. Some of these Stake has identified as "issues." His "issues" exist when there is a factor that the participants agree is important to the situation but about the consequences of which they disagree. In the rather atomistic symbolism we have used before: when an X is agreed by the participants to be important and separately they say its affect is L or M or N or O,... then this is an "issue" of the situation. Case study research sets out (among other things) to describe "issues." The issue belongs to the participants, it is not invented by the researcher, its description is in their terms and they are responsible (not the researcher) for its validation.

The end of the second pathway is not, of course, the whole reality of the original situation but the path of the research has been one of trying to capture, as well as words (or pictures etc.) can, the fullness of some aspects of the reality the participants experience while in it. The process is one of trying to avoid abstraction. The reality the researcher seeks, at least for Stake (1979), is the reality of the situations - nothing more. Indeed, he seems to suggest that research will always fall short of reality and certainly not extend it as traditional researchers aim to do.

A number of fundamental differences between the two research approaches are I think obvious and sharply differentiate the commitments the two groups of researchers hold. Only one will be further mentioned since it was to me a quite new insight. Traditional research is governed by $O_i - O_j$ relationships across a population. The very nature of this process ensures that it will not throw up the "issues" of the other approach which are non-consistent relationships - lost in statistical insignificance.

IV. Paradigm is the sharing of examples.

The discussion above (in III) helps to differentiate what examples are shared in the two approaches. Traditional researchers can see commonality in a variety of studies of say teachers' style. Each of them can be reduced to a study of the relations between teacher properties defined by the researcher and other variables in the situation such as certain student

learning outcomes. They are all examples of an O_i being related to an O_j. The examples are particular to each study, but they belong to a larger class designated by a still more inclusive pair of variables. Their "explanation" are examples of a general atomistic view of the nature of educational situation - the "wood chip" view.

Case study workers share descriptions of "issues" in a similar manner. They see them as examples of the same sort of reality. "Issues" are to them important ingredients (probably more so than agreed generalisations) than can provide the creative (or disruptive) energy of the educational process. They see them as having the power, that examples in Kuhn's paradigms have, although it is not a predictive power as in the traditional approach. The power of good descriptions of issues lies, they believe, in the fact that they can assist the understanding of the participants and perhaps inform other participants in similar (but never identical) social situations. Increased understanding of this sort leads to more informed (though still subjective) judgments for future action (see Stenhouse, 1978; and Stake, 1978). The sharing in this approach thus naturally extends beyond the researchers to the participants who have provided the issues and validated their descriptions.

In considering the shared examples of the two groups, it is important to get beyond some superficial confusions of language. For example, traditional researchers can and do study single cases. They would do this as precursor to, or part of a wider study. Their studies of the single case, if they involve some steps of the upper path in Figure 1, do not provide examples that "case study" researchers would share. Again, a case study researcher invariably works in the field but so often do survey analysts. That both do "field work" again does not mean that they would recognise as co-examples what each other does while in the field.

A sharp participant in my own early case study told me one day that he had just learnt the difference between an industrial psychologist (a survey analyst who had spent the morning collecting data) and a researcher (myself):

> An industrial psychologist tells you what is important about your situation, while a researcher asks you what is important and waits for the answer.

SUMMARY

For each of the four distinctive aspects of a paradigm, case study research does now seem to stand sharply apart from traditional research. It seems then, that it is appropriate to speak of "competing paradigms" and that we in science educational research would do well to examine our own practices as researchers and see on which side we wish to stand.

REFERENCES

CAMPBELL, D.I. III "Degrees of freedom" and the case study, Comparative Political Studies, 1975, 8, 178-193.

CRONBACH, L.J. Beyond the two disciplines of scientific psychology, American Psychologist, 1975, 30, 116-127.

ELLIOT, J. Ford Teaching Project Reports. C.A.R.E., University of East Anglia, 1973-75.

ERAUT, M. In Draft Chapter 1 of a new book, p.7, University of Sussex, 1978.

FENSHAM, P.J. & HOOPER, D. The Dynamics of a Changing Technology. London: Tavistock Press, 1959.

FENSHAM, P.J. & INGVARSON, L.C. Two pairs of eyes are better than one, to be published in Journal of Curriculum Studies, 1979.

GARDNER, P.L. Research on teacher effects: critique of a traditional paradigm, British Journal of Educational Psychology, 1974, 44, 123-130.

HAMILTON, D. Handling innovation in the classroom: two Scottish examples. In Reid, W.A. & Waller, D.F. (eds.) Case Studies in Curriculum Change. London: Routledge and Kegan Paul, 1975.

HAMILTON, D. Some contrasting assumptions about survey analysis and case study research. Paper privately circulated, University of Glasgow, 1978.

ITTELSON, W.H. & CANTRILL, H. Perception - A Transactional Approach. New York: Doubleday Papers in Psychology, 1954.

KUHN, T.S. The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1962.

KUHN, T.S. Postscript - 1969. In the 2nd Edition of The Structure of Scientific Revolutions. Chicago: University of Chicago Press, 1970.

MacDONALD, B. & WALKER, R. Case study and the social philosophy of educational research, Cambridge Journal of Education, 1975, 5, 2-11.

MASTERMAN, M. The Nature of a Paradigm. In Lakatos, I. & Musgrave, A. (eds.) Criticism and the Growth of Knowledge. Cambridge: Cambridge University Press, 1970.

NICODEMUS, R.B., JENKINS, E.W. & INGLE, R.B. Adopting Nuffield O- and A-level Chemistry, Education in Chemistry, 1976, 13, 46-48.

OWEN, J.M. & TISHER, R.P. Curriculum adoption: the fate of a national curriculum project in Australia. In Proceedings of Bat Sheva Conference, Science Teaching Centre, Hebrew University, Israel, 1979.

PARLETT, M. & HAMILTON, D.A. Evaluation as illumination: a new approach to the study of innovatory programmes. Reproduced in Hamilton et al. (eds.), Beyond the Numbers Game. London: Macmillan Education, 1977.

PARSONS, C. The New Evaluation: A cautionary note, Journal of Curriculum Studies, 1976, 8, 125-138.

PHILLIPS, D.C. Conservative influences on experimental research, Research in Science Education, 1974, 5-10.

POWER, C.N. Competing paradigms in science education research, Journal of Research in Science Teaching, 1976, 13, 579-587.

ROTHMAN, A.I., WALBERG, H.J. & WELCH, W.W. Effects of a summer institute on physics teachers, Science Education, 1968, 52, 469-473.

ROSIER, M.J. Implications of the results of the IEA Science Project for the teaching of science in Australia, Research in Science Education, 1974, 219-227.

SNOW, R.E. Representative and quasi-representative designs for research on teaching, Review of Educational Research, 1974, 44(3), 265-292.

STAKE, R. The case study method in social inquiry, Educational Researcher. 1978, 1, 5-8.

STAKE, R. What social science does for education. Privately circulated statement, C.I.R.C.E., University of Illinois, 1979.

STAKE, R. & EASLEY, J. et al. Case studies in science education. Washington: Government Printing Office, 1978.

STENHOUSE, L. An introduction to curriculum research and development. London: Heineman, 1975.

STENHOUSE, L. Case study and case records: towards a contemporary history of education, Journal of British Education Research Association, 1978.

SUAN, M.Z. Evaluation of A.S.E.P.: a case study approach. M.Ed. thesis, Monash University, Australia, 1976.

WALKER, R. Studies on Nuffield Secondary Science as part of SAFARI, C.A.R.E., University of East Anglia, 1974.

ZANGWILL, O.L. An introduction to modern psychology. London: Methuen, 1950.

Research in Science Education,1979,9,53-54

THE AIMS OF SCIENCE COURSES

N. R. Thorley, D. J. Boud, J. C. Dunn and T. Kennedy

INTRODUCTION

A project at the Western Australian Institute of Technology (WAIT) has been concerned with what science courses, and laboratory courses in particular are, or should be, achieving (Boud et al., 1978). Opinion has been sought from key groups on two main issues:

(i) what views undergraduate students have about the general aims of science courses, and how these compare with similar studies in other countries.

The general aims questionnaire used was that developed by Frazer, Palmer and Sleet (1976). Science undergraduate students were asked to rate each of thirty three aims in terms of "observed" and "preferred" importance in their courses. Responses were analysed by calculating for each aim the mean importance rating. Comparisons were then made between different survey groups' ranking of the aims:

(a) on the "preferred importance" scale, and

(b) on a "discordance" scale, obtained by subtracting "observed" from "preferred" importance.

Agreement between year and subject groups within the local sample was close. The aims most highly rated on the "preferred importance" scale were basic cognitive and attitudinal ones, whilst those lowly rated tended to be associated with social and economic issues. Reasonable overall agreement was obtained with the findings of studies in the U.K. (Frazer, Palmer & Sleet, 1976) and South Africa (Lynch & Gerrans, 1977). More interesting to course designers are the "discordant" aims. Again the local and U.K. results were similar. Of the eleven most discordant aims in each survey, seven were common. Top in both lists were "application to environmental problems," and others related to higher level intellectual skills.

(ii) what views practising scientists, recent graduates and undergraduates have about the aims of laboratory courses.

The laboratory aims questionnaire (Boud, 1973) consisted of twenty two items. Students were asked to rate aims in terms of importance both on a "preferred" and "observed" scale, but those in employment were only requested to rate on a "preferred" scale. A high degree of concordance was found between the practising scientist group (which included tertiary teachers) and recent graduates. These groups regard laboratory work in pragmatic terms, which stress the acquisition of various skills such as

observation and interpretation of data. Students, however, have different perceptions of laboratory aims, and rate highly those activities associated with educational processes, for example the links which they observe to exist between taught theoretical material and laboratory activities.

DISCUSSION

There is a broad general agreement between students in the U.K., South Africa and Australia on the relative importance that they felt should be accorded each general aim. There are also some striking points of agreement in the deficiencies they see in their courses.

However the views of students and those in employment disagree about laboratory aims. This raises a crucial issue; what role should student opinion have in course design. Although employers, graduates and teaching staff are in the better position to judge the immediate needs of industry, it is probably fair to suggest that the ultimate qualities of a graduate are as much determined by his attitude to the course as by the content of it.

We offer the following as suggestions, realising that not all of these may be suitable in any one situation: the views of scientists, graduates and employers could be presented to students so that they may gain some perspective on their own opinions; the design and presentation of courses could be made in such a way that the inputs of practitioners and industry were clearly indicated; or perhaps consultative groups be established to discuss course offerings and include students, tertiary staff, recent graduates and practitioners. Whatever the solution, the challenge remains that of the staff who will conduct the course.

REFERENCES

BOUD, D.J. The laboratory aims questionnaire: A new method for course improvement. Higher Education, 1973, 2, 23-26.

BOUD, D.J., DUNN, J.G., KENNEDY, T. & WALKER, K.G. Laboratory teaching in tertiary science: A review of some recent developments. Higher Education Research and Development Society of Australasia, University of New South Wales, 1978.

FRAZER, M.J., PALMER, C.R. & SLEET, R.J. Aims of first degree courses: Student opinion. Education in Chemistry, 1976, 13, 44.

LYNCH, P.P. & GERRANS, G.C. Aims of first year courses in chemistry, according to staff, and students' expectations of such courses. South African Journal of Science, 1977, 73, 25-27.

Research in Science Education,1979,9,55-68

STUDENTS' ATTITUDES TO AN INNER-CITY ENVIRONMENT

W. D. Ponniah

INTRODUCTION

This paper is in two parts. The first part describes the development and validation of a modified Semantic Differential instrument for measuring attitudes toward the inner-city environment. The second part compares the differences in attitudes between two samples of school children (one from the inner-city and the other from the outer suburbs) to determine if any change in attitude occurred as a result of a two week long instructional programme.

DEVELOPMENT OF THE SEMANTIC DIFFERENTIAL INSTRUMENT

Most published environmental attitude studies have used questionnaires, interviews, checklists or summated rating scales (Likert-type) as the instrument of attitude measurement; the Semantic Differential technique has not been as popular. The Semantic Differential was, however, selected for this study because of its greater potential for reflecting the complexity of multidimensional concepts than other traditional techniques which have been in use for the last two decades.

Some modification in the method presented by Osgood (1957) was used in this study. The first modification was in the selection of the following factors - Beautification, Optimism and Pollution (BOP) in place of the more familiar Evaluative, Potency and Activity (EPA) factors. The second modification was in the mode of presentation of concept stimuli. In place of the usual written word, projected 35mm colour slides were used.

Richmond in Victoria was chosen for this study because it possesses most of the characteristics of a typical inner-city environment. It has areas devoted to high density housing, narrow congested streets and roads, miles of drab concrete factory walls as well as concrete freeways encroaching into valuable parklands. In addition Richmond has one of the most polluted stretches of the Yarra River flowing through it. Of further interest is the fact that Richmond has a few remaining wilderness areas which serve as reservoirs for native fauna and flora as well as a few remaining riverside recreational facilities for boating, golfing, and picnicing.

After visits to the area and discussion with fellow researchers, the following six concepts were chosen: Yarra, Richmond, Freeway, Factory, Wilderness, Recreation.

Attitudes to be Assessed.

The attitudes to the above concepts which I wished to assess were:

(a) realisation of one's own worth;

(b) respect for others;

(c) aesthetic appreciation of surroundings;

(d) hopefulness for the future;

(e) responsibility.

As it was considered difficult to assess the above attitudes directly, it was decided to concentrate on the following three dimensions:

(a) Beautification dimension: which refers to the qualities of the inner-city environment which affords keen pleasure to the senses;

(b) Optimism dimension: which refers to the degree of optimistic feelings (hopefulness) for the altered inner-city environment;

(c) Pollution dimension: which refers to the degree of pollution (air, water, noise, visual) of the inner-city environment.

Item Preparation and Selection.

The next step in the development of the instrument was the listing of suitable sets of bipolar adjectival scales for each of the factors B, O, and P. Copies of the scales were then distributed to English teachers in forms one, two, and three in Mount Scopus College, Burwood, as well as to researchers and lecturers in the Faculty of Education, Monash University, Clayton, in order to seek consensus on the readability of each item for a target population of form three students.

On the basis of feedback from teachers, fellow researchers and lecturers, difficult items were omitted from the scales whilst minor changes were made to a few others.

From the 300 slides taken, 18 slides (3 per each concept category) were chosen on the basis that they best depicted the environmental scenes in each of the concept categories. Ten items from each list (B, O, & P) were chosen to suit each of the 18 slides. This resulted in the design of 3 booklets each of 180, five-step items which will henceforth be referred to as the B, O, and P Environment Attitude Scales.

FIELD VALIDATION PROCESS

A pilot study was conducted in October 1977 with the aim of selecting relevant test-items from each of the B, O, and P scales so as to reduce the 3 sets of 18 scales and 180 items to a single set of 18 scales and 180 items. The three scales were field tested with 125 form two, three, and

four students from Collingwood and Prahran High Schools. Every first, fourth, seventh, etc., child was given the B booklet; every second, fifth, eighth, etc., child was given the O booklet; every third, sixth, ninth, etc., child was given the P booklet. They were told to work through the list as each of the 18 slides were projected by means of a Carousel on to a screen.

Mean scores were obtained for each subject on each of the three dimensions for each concept. As there were 3 factors and (3 x 6) concepts evaluated, each subject generated 54 mean scores on each of the pre- and post-test measures.

The scores on the tests were subjected to item analysis. Each of the 18 concepts in each of the B, O and P scales were treated as separate scales and the correlations with scores of the ten B, O, and P items on their respective scales were considered. The three best B and P items and the four best O items in each scale were selected to make up the 18 new scales of ten items per concept category.

The semantic differential finally used contained the following scales and subscales:

CONCEPT	SUBSCALES		CONCEPT	SUBSCALES	
Yarra (YA)	Beautification	YA-B	Factory (FA)	Beautification	FA-B
	Optimism	YA-O		Optimism	FA-O
	Pollution	YA-P		Pollution	FO-P
Richmond (RI)	Beautification	RI-B	Wilderness (WI)	Beautification	WI-B
	Optimism	RI-O		Optimism	WI-O
	Pollution	RI-P		Pollution	WI-P
Freeway (FR)	Beautification	FR-B	Recreation (RE)	Beautification	RE-B
	Optimism	FR-O		Optimism	RE-O
	Pollution	FR-P		Pollution	RE-P

MAIN STUDY

The study involved the measurement of environmental attitudes of two samples of form three students from Richmomd Girls' High School (Richmomd) and Cleeland High School (Dandenong) by administering the B, O, and P scales. Some simple and contrived observations of overt behaviour were also obtained. The design used is shown below:

RICHMOND GIRLS' HIGH SCHOOL	Experimental Group	R	O_1	X	O_2
		R		X	O_3
	Control Group	R	O_4		O_5
		R			O_6

CLEELAND HIGH SCHOOL	Experimental Group	R R	O_1	X X	O_2 O_3
	Control Group	R R	O_4		O_5 O_6

where O_1, O_4 = pretests

O_2, O_3, O_5, O_6 = posttests

X = learning experience

R = randomised subjects

This design has much in common with the design which Campbell and Stanley (1966) refer to as the Solomon Four-Group Design.

In all, 89 students from four form three classes (two each from Richmond Girls' High School and Cleeland High School) took part in the study which was conducted during the last two weeks of November 1977. The sample consisted of 42 students from RGHS and 47 students from CHS. In both schools intact classes were used for treatment and control groups, but within each group the pupils were randomly assigned to the pre- and post-tests or post-test only groups.

Programme for the Experimental Group.

The instruction provided consisted of a multi-media presentation of colour slides correlated with audio-tapes of background sound effects. There were six lessons of approximately 35 minutes' duration each. One lesson was based on each of the six concepts and the lessons were of the slide discussion type with discussions based upon the chosen slides. These lessons were not propaganda, that is, they were not one-sided or heavily loaded on the disadvantages of inner-city living. Recordings were played at the appropriate times as background sound effects.

Tests.

The pre- and post-tests were similar in design and were based on the modified Semantic Differential instrument developed and revised by the researcher. The pre-tests were administered during the first class session of the unit whilst the post-tests were administered during the last class session.

Unobtrusive Behaviour Measures.

In addition to the formal tests (pre and post) two other unobtrusive measures were employed in the study in an attempt to focus on behavioural outcomes. These were:

(a) optional questionnaires placed on a bench in the class;

(b) newspaper articles on current environmental issues pinned on the class bulletin board.

Before the two sets of questionnaires were placed on the bench the students were informed that it was not compulsory for them to take a copy of the questionnaire but that should they be interested they could take their copy home to discuss with family members and friends. They were also informed that if they took a copy home they could return the completed forms to their class teachers the following day.

For the second unobtrusive measure, I announced to the classes concerned that some newspaper articles on current environmental issues were on the class bulletin board and that should they come across any interesting articles on any environmental issue they could pin them up alongside the others.

In both cases, the number of students responding and the nature of the responses were carefully noted by the researcher as well as by their respective classroom teachers.

Following the administration of the formal test instrument, data were tabulated for pre- and post-test results.

RESULTS

Repeated Measures ANOVA of Post-Test Scores.

This section concerns the repeated measures ANOVA of the post-test data of 46 subjects of the combined pre-test - post-test group which was carried out according to the method described by Winer (1971). This is a factorial analysis where the same experimental unit or subject is observed under more than one condition. The aim in carrying out repeated measures ANOVA was to test the research hypotheses listed under Tables 1-7 appended to this paper.

Summary of Findings of Repeated Measures ANOVA.

Tables 1-7 were prepared by considering each one of the following seven sources in turn and summarising the results on each source:

(a) location of school (Richmond vs. Cleeland);

(b) type of treatment (experimental vs. control);

(c) 2-way interaction between school and treatment;

(d) 2-week long interval between pre- and post-tests (time factor);

(e) 2-way interaction between location of school and time interval;

(f) 2-way interaction between treatment and time interval;

(g) 3-way interaction between location of school, treatment and time interval.

Table 1 clearly indicates that children in a school located in the outer suburbs displayed different attitudes on eight sub-scales (RI-B, RE-O, RI-P, FR-B, FA-B, FA-O, FA-P, WI-B) to those of children in a school located in the inner-city. As this result was contrary to what was predicted in Hypothesis 1, the latter was therefore rejected.

The combined control group (Richmond and Cleeland) displayed different attitudes on only one of the sub-scales (FR-B) to those in the combined experimental group (Richmond and Cleeland) (Table 2). There is insufficient information in one significant difference on 18 sub-scales to reject Hypothesis 2.

Table 3 shows that the control group students from both schools displayed different attitudes on two sub-scales (FR-B, FR-O) to those of the experimental group students. As the prediction of Hypothesis 3 (a, b) was contrary to the results obtained above, Hypothesis 3 was rejected.

Table 4 demonstrates that an interval of two weeks between the administration of the pre- and post-tests did affect results on the post-test on eight sub-scales (YA-B, YA-O, RI-B, FR-B, FR-O, FR-P, FA-O, WI-P). Hence, Hypothesis 4 rejected as its prediction was contrary to the result obtained.

Hypothesis 5 predicted that students in a school located in the outer suburbs will not display different attitudes to any of the 18 sub-scales over time as those in a school located in the inner-city.

But the 2-way interaction between school and time-interval did affect the post-test results on the FR-O, FA-B and FA-P sub-scales, (Table 5). As the prediction of Hypothesis 5 was contrary to the result obtained, Hypothesis 5 was rejected. It appears that the Freeway controversy which raged at the time of the experiment had a different impact on students in the two locations.

The 2-way interaction between treatment and time (Table 6) did affect results on the FA-B, FA-P and WI-B post-test scores. As this result was contrary to what was predicted, Hypothesis 6 was therefore rejected. Treatment is associated with the development of more positive attitudes to the Factory Beautification and Factory Pollution and more negative attitudes to Wilderness Beautification, but there was no evidence of an impact on any of the other 15 dimensions.

Table 7 deals with the 3-way interaction between school, treatment and time interval. The postulated Hypothesis 7, which was in four parts, predicted that students in the experimental and control groups in the inner-city and outer suburban schools would not display different attitudes to any of the 18 sub-scales over time.

But the results indicated that such an interaction occurred for attitudes on the FR-O, FR-P, and WI-B sub-scales. As this result was contrary to what was predicted, Hypothesis 7 was rejected as well.

CONCLUSION

In analysing the results of Tables 1-7, one is led to the conclusion that if environmental education is to have a significant impact in our schools, it has to have some educational relevance, that is, that it must fit into the real life context of the child.

This study confirmed that children in an inner-city school had more positive attitudes about an inner-city environment than the outer suburban children who were more concerned with the aesthetically pleasing aspects of wilderness areas (Tables 1 and 7). It seems that the further the reference was from the respondents' immediate surroundings, the less reluctant the outer suburban students were to identify significant problems like pollution.

Only by taking into consideration such basic information about each group of children - inner-city, outer suburban or rural - can environmental education programmes be planned so as to present to each group of students the most relevant and effective programme.

REFERENCES

CAMPBELL, D.T. & STANLEY, J.C. Experimental and quasi experimental designs for research. Chicago: Rand McNally and Co., 1966.

OSGOOD, C.E., SUCI, G.J. & TANNENBAUM, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1959.

WINER, B.J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

TABLE 1

Direction of Differences for School as a Source of Variance.

No.		Sub-scale	F-value	Mean Scores in cases of Significant F-values RICHMOND	CLEELAND
1	YARRA (YA)	-B	0.24		
		O	0.47		
		P	0.78		
2	RICHMOND (RI)	-B	35.23 **	23.98	17.69
		O	31.67 **	44.86	35.90
		P	11.14 **	25.82	21.81
3	FREEWAY (FR)	-B	14.44 **	28.41	22.48
		O	2.42		
		P	2.19		
4	FACTORY (FA)	-B	16.96 **	13.98	10.31
		O	14.98 **	30.59	23.25
		P	7.34 **	16.09	13.48
5	WILDERNESS (WI)	-B	5.54 *	43.59	44.71
		O	0.94		
		P	1.31		
6	RECREATION (RE)	-B	0.28		
		O	0.18		
		P	1.64		

* $p < 0.05$ ** $p < 0.01$

Research Hypothesis

1. That students in a school located in the outer suburbs will not display different attitudes to any of the 18 sub-scales to those of students in a school located in the inner-city.

TABLE 2

Direction of Differences for Treatment as a Source of Variance.

No.		Sub-scale	F-value	Mean scores in cases of significant F-values EXPERIMENTAL	CONTROL
1	YARRA (YA)	-B	0.47		
		O	0.16		
		P	1.89		
2	RICHMOND (RI)	-B	0.25		
		O	0.07		
		P	3.46		
3	FREEWAY (FR)	-B	4.68*	23.76	27.13
		O	0.03		
		P	0.77		
4	FACTORY (FA)	-B	0.15		
		O	0.00		
		P	0.04		
5	WILDERNESS (WI)	-B	0.02		
		O	0.02		
		P	0.06		
6	RECREATION (RE)	-B	0.59		
		O	1.45		
		P	0.00		

* $p < 0.05$ ** $p < 0.01$

Research Hypothesis

2. That students in the combined control group (Richmond and Cleeland will not display different attitudes to any of the 18 sub-scales as those in the combined experimental group (Richmond and Cleeland).

TABLE 3

Direction of Differences for School X Treatment as a Source of Variance

Scale	Sub-Scale	F-values	Mean Scores in cases of significant F-values			
			RICHMOND		CLEELAND	
			Experimental	Control	Experimental	Control
1 YARRA (YA)	-B	1.65				
	O	0.92				
	P	0.01				
2 RICHMOND (RI)	-B	0.05				
	O	0.07				
	P	0.51				
3 FREEWAY (FR)	-B	7.14*	24.64	32.18	22.88	22.08
	O	10.46**	48.77	54.91	51.79	46.29
	P	1.19				
4 FACTORY (FA)	-B	1.19				
	O	2.16				
	P	3.89				
5 WILDERNESS (WI)	-B	0.69				
	O	1.51				
	P	1.46				
6 RECREATION (RE)	-B	0.159				
	O	0.14				
	P	2.92				

* $p < 0.05$ ** $p < 0.01$

Research Hypothesis

3(a) That the control group students in the inner-city school will not display different attitudes to any of the 18 sub-scales to those in the experimental group.

3(b) That the control group students in the outer suburban school will not display different attitudes to any of the 18 sub-scales to those in the experimental group.

TABLE 4

Direction of Differences for Time Interval as a Source of Variance.

	Scale	Sub-scale	F-value	Mean Scores in cases of significant F-value	
				PRETEST	POSTTEST
1	YARRA (YA)	-B	17.30**	11.91	14.86
		O	14.96**	18.60	23.55
		P	1.01		
2	RICHMOND (RI)	-B	4.96*	20.23	21.44
		O	1.04		
		P	1.01		
3	FREEWAY (FR)	-B	16.67**	23.85	27.04
		O	10.48**	48.23	52.66
		P	4.46*	24.84	27.29
4	FACTORY (FA)	-B	2.97	24.67	29.17
		O	11.36**		
		P	3.10		
5	WILDERNESS (WI)	-B	0.06		
		O	2.34		
		P	6.12*	43.88	42.92
6	RECREATION (RE)	-B	0.37		
		O	3.33		
		P	3.20		

* $p < 0.05$ ** $p < 0.01$

Research Hypothesis

4. That an interval of 2 weeks between the administration of the pre- and post-tests will not affect results on the posttest.

TABLE 5

Direction of Differences for School x Time Interval as a Source of Variance.

	Scale	Sub-Scale	F-value	Mean Scores in cases of significant F-values			
				RICHMOND		CLEELAND	
				Pretest	Posttest	Pretest	Posttest
1	YARRA (YA)	-B	0.34				
		O	0.28				
		P	0.68				
2	RICHMOND (RI)	-B	0.22				
		O	2.03				
		P	0.34				
3	FREEWAY (FR)	-B	0.06				
		O	7.14*	51.45	52.23	45.00	53.08
		P	0.15				
4	FACTORY (FA)	-B	8.74**	12.32	15.64	10.75	9.88
		O	3.05				
		P	8.49**	14.32	17.86	13.92	13.04
5	WILDERNESS (WI)	-B	0.62				
		O	3.21				
		P	0.05				
6	RECREATION (RE)	-B	0.37				
		O	0.29				
		P	0.06				

* $p < 0.05$ ** $p < 0.01$

Research Hypothesis

5. That students in a school located in the outer suburbs will not display different attitudes to any of the 18 sub-scales over time to those in a school located in the inner city.

TABLE 6

Direction of Differences for Treatment x Time-Interval as a Source of Variance.

Scale	Sub-Scale	F-value	Mean Scores in cases of significant F-values			
			EXPERIMENTAL		CONTROL	
			Pretest	Posttest	Pretest	Posttest
1 YARRA (YA)	-B	2.78				
	O	0.38				
	P	0.04				
2 RICHMOND (RI)	-B	0.09				
	O	0.19				
	P	0.32				
3 FREEWAY (FR)	-B	1.43				
	O	1.42				
	P	0.58				
4 FACTORY (FA)	-B	5.16*	10.90	13.73	12.17	11.78
	O	0.53				
	P	4.24	13.43	16.33	14.80	14.58
5 WILDERNESS (WI)	-B	11.71**	44.41	43.83	43.84	44.52
	O	1.34				
	P	0.89				
6 RECREATION (RE)	-B	0.80				
	O	1.84				
	P	0.03				

* $p < 0.05$ ** $p < 0.01$

Research Hypothesis

6. That students in the combined control group (Richmond and Cleeland) will not display different attitudes to any of the 18 sub-scales over time as those in the combined experimental group (Richmond and Cleeland).

TABLE 7

Direction of Differences for School X Treatment X Time Interval as a Source of Variance

Scale	Sub-Scale	F-values	Mean Scores in cases of significant F-values							
			RICHMOND				CLEELAND			
			Experimental		Control		Experimental		Control	
			Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
1 YARRA (YA)	-B	0.15								
	O	0.27								
	P	0.10								
2 RICHMOND (RI)	-B	0.30								
	O	0.51								
	P	0.12								
3 FREEWAY (FR)	-B	1.72								
	O	8.71**	47.18	50.36	55.73	54.09	50.88	53.0	39.42	53.17
	P	4.36	23.36	27.82	28.73	30.09	25.50	24.17	21.75	27.08
4 FACTORY (FA)	-B	2.28	43.82	42.91	43.27	44.36	45.00	44.75	44.42	44.67
	O	3.22								
	P	0.33								
5 WILDERNESS	-B	4.22*								
	O	0.24								
	P	1.10								
6 RECREATION	-B	1.21								
	O	1.33								
	P	3.51								

* $p < 0.05$ ** $p < 0.01$

Research Hypothesis

7(a) That students in the experimental group in an inner-city school will not display different attitudes to any of the 18 sub-scales over time.

7(b) That students in the control group in an inner-city school will not display different attitudes to any of the 18 sub-scales over time.

7(c) That students in the experimental group in an outer-suburban school will not display different attitudes to any of the 18 sub-scales over time.

7(d) That students in the control group in an outer-suburban school will not display different attitudes to any of the 18 sub-scales over time.

68

Research in Science Education,1979,9,69-74

THE RELATIONSHIP BETWEEN ATTITUDES AND SCHOOL ACHIEVEMENT: A FURTHER EXPLORATION

M. N. Maddock

THE IMPORTANCE OF ATTITUDE ASSESSMENT

In the past few years there has been a steadily increasing emphasis on school-based curriculum development, school-based assessment and evaluation and so the importance of attitudinal aims in the curriculum. Science teaching has been no exception in this trend (Thomson, 1975; Gardner, 1975; Sullivan, 1975).

The Secondary Schools Board of the N.S.W. Department of Education has taken attitudinal aims quite seriously, giving clear specifications for this area in its aims document for science (Secondary Schools Board, 1973 a) and in a policy statement that attitudes and interests are to be assessed stated in the assessment guide (Secondary Schools Board, 1973 b). This position has placed N.S.W. secondary science teachers in a position of considerable dilemna, as indicated by McBurney's evaluation study (McBurney, 1978).

Developing out of an interest in the evaluation of attitudinal aims in science curricula, which resulted in a major research study in Papua New Guinea (Maddock 1977 a, Maddock 1977 b), the writer commenced an evaluation study of attitudinal aims in the N.S.W. Science Syllabus in 1977 (Maddock, 1978).

The writer has been particularly concerned about the development of evaluative instruments for attitudinal aims and the relationship between scores on attitude scales and scores on achievement scales, because of the N.S.W. Secondary Schools Board policy on the additive use of knowledge and skills achievement test results with attitude assessment (Secondary Schools Board, 1973 b). The preliminary exploratory study in 1977 (Maddock, 1978) developed a scale to evaluate the objective:

> to develop a belief in the value of the conservation of our natural environment (Secondary Schools Board 1973 a, p. 11).

In 1978, this study was repeated and the scope extended to another attitude area, that of "critical thinking" or "open-mindedness" as expressed in the syllabus aim.

> 1.2 To develop a belief in the value of:
>
> 1.21 Thinking critically in an unbiased and open-minded manner (Secondary Schools Board, 1973 a, p. 9).

INSTRUMENTATION

The Environmental Conservation Attitude Scale.

Based on the item analysis results obtained in 1977, the 22 items of the pilot scale were revised and an additional item added. The revised instrument was trialled with a sample of 1,120 subjects from five high schools in the Newcastle area (grades 7-11) and a first year class at the Newcastle College of Advanced Education. A further trial was conducted with grade 10 students at a country high school, where the school programme had an environmental emphasis. Analyses were based on the revised original 22 items, for comparison with the 1977 study, and on the total instrument, including the new, 23rd item.

Critical-Thinking or Open-Mindedness.

Using a similar technique of specifying the expression of an attitude in behavioural terms, the construction of the item stem by describing a concrete situation and asking for a choice from three alternative specific behaviours which could result from the situation described, as was done in the Papua New Guinea study (Maddock, 1977 b) and in the development of the environmental scale (Maddock, 1978), a pool of items for the "critical thinking" attitude scale was developed. This pool was then submitted to an editing committee which finalised a 21 item pilot version. This was then trialled in the same five schools as the environmental scale, using a sample of 312 pupils in grades 7, 8, 10 and 11, in classes which had also been used for the environmental attitude scale.

Classroom Achievement Tests.

As was done in the 1977 study, normal classroom achievement tests were set and administered to a sample of subjects who had done both the environmental conservation attitude and the "critical thinking" scales. The two attitude scales were correlated together across the total sample and three way correlations were done between these two scales and the separate achievement tests. In addition, one way analyses of variance across grade levels were carried out on the attitude scales.

In the case of the environmental conservation attitude scale, the across grade analysis was carried out using the 22 item version, in order to make comparisons with the findings in 1977.

RESULTS

Environmental Conservation Attitude Scale.

The environmental scale scores demonstrated a much improved level of reliability over the 1977 pilot version, the 22 item version producing a coefficient alpha of 0.65, the 23 item version 0.66, taken over the total sample, and the 23 items gave an alpha of 0.72 on the additional grade 10

sample at the country high school which was pursuing an environmentally oriented programme. Alphas of this magnitude indicate that the results could be used for useful comparisons across groups.

Critical Thinking Scale.

The 21 item pilot version of the "critical thinking" scale produced an encouraging coefficient alpha of 0.56, which indicates that with further revision and possibly an increased number of items it can be developed into a viable scale. The alpha was considered high enough for a preliminary cross-grade analysis to compare means. The alphas obtained at the various schools ranged between 0.38 and 0.70.

Science Achievement Tests.

The range of reliability coefficients was quite wide, ranging from 0.31 to 0.80, but seven of the achievement tests had coefficient alpha reliabilities above 0.5, making them useful for considering correlations between the three scales. These results were very similar to those found in 1977 (Maddock, 1978).

Correlations Between Scales.

In one school, the same achievement test was used on three grade 7 classes, a total of 86 students. In all other cases, separate achievement tests were used. The correlations obtained in 1977 between achievement tests and the attitude to environmental conservation were all low (0.10). A similar pattern was found in the 1978 study for each class. For the combined classes the correlation was 0.12.

In the case of achievement - critical thinking, correlations were either negative or non-significantly low and positive, ranging from - 0.47 to + 0.32 for individual classes, and 0.27 for combined classes.

Corrrelations between the conservation scale and the critical thinking scale were also low. Correlations in combined classes in two schools were + 0.22 and + 0.31 respectively, while the correlation for the combined sample of all schools was + 0.37 (significant at 0.01).

It can be seen that although there is a fair amount of variation in the correlations obtained, the correlations between the attitude to conservation of the environnent scale, the "critical thinking" scale, and science achievement tests, are in general low, indicating that the measures are relatively unrelated, measuring different attributes.

Analyses of Variance by Grade Level.

The one way analysis of variance across grade levels for the 22 item version of the scale gave a significant F ratio of 4.85 ($p < 0.001$). The grade 7 mean was almost identical to that obtained in 1977. The grade 8 mean was 2.1 scale points higher than that obtained in 1977, but was almost

identical to the 1977 grade 7 mean and the means for grades 9 and 11 were almost identical to the grades 8 and 9 means for 1977.

Thus there is no significant difference between the means of grades 8, 9 and 11 in 1978 and the means of grades 7, 8 and 10 in 1977.

In 1977 it was the difference between the combined grades 9-10 and the combined grades 7-8 which was significant, as determined by the Scheffe test (Kerlinger and Pedhazur, 1973, p. 128-131). In 1978, the difference of 1.4 points between these two was again significant, but this time in favour of the combined grades 7-8. However, the actual difference is a very minor one in an educational sense, and supports the contention that the effect of grade level is minimal.

The results are summarised in Table 1.

TABLE 1

Comparison of Mean Attitude to Conservation across Grades.

Grade	1977	1978
CAE	-	51.6
11	52.6	51.0
10	51.2	49.6
9	-	48.6
8	48.5	50.6
7	50.5	50.7

The across-grade, one way analysis of variance for the "critical thinking" attitude scale produced an F ratio of 9.02 (df 3/307), significant at the 0.001 level. Inspection of the grade means (see Table 2) shows that the significant difference occurs between grade 11 and grade 10, the other means being very close.

TABLE 2

ANOVA by Grade Level, Critical Thinking Scale.

Grade	Mean	ANOVA	
11	54.5	SS	518.43
10	49.7	F	9.02
8	50.8	df	3/307
7	50.2	p	< 0.001

CONCLUSIONS

The study provides further evidence in support of the usefulness of using behavioural specifications for concrete situations as a technique for developing reliable attitude scales, a technique which was pioneered in the Papua New Guinea studies (Maddock, 1977).

The 1978 revised version of the environmental conservation attitude scales produced a much improved level of coefficient alpha reliability over the 1977 pilot version, to a stage when the instrument can be used for group comparisons. A pilot form of a new scale, a critical thinking - open mindedness attitude scale, also gave a satisfactory reliability coefficient, indicating that it has potential for further development and improvement.

The results of correlational analyses support the findings in the 1977 study that there is little relationship between the attitude towards the conservation of the environment, as measured by the attitude scale, and achievement based on classroom tests in science. Correlations were also low between the "critical thinking" scale results and achievement test results, thus suggesting that all three are independent measures addressing different attributes. This throws further doubt on the validity of using such attitude measures with achievement tests for additive norm-referenced assessment in schools, as suggested by N.S.W. Secondary Schools Board policy.

The results of across-grade analyses with both the environmental and the "critical thinking" scale raise doubts as to whether secondary schooling, with its science teaching component, is having any effect on achieving the stated attitudinal aims in the N.S.W. science syllabus. Although many variables are involved which were not controlled for in the studies, it would be expected that if science teaching was to be considered as having any significant shift with increasing grade level ought to be detected.

As in 1977, the environmental scale showed a statistically significant difference across grade levels but not one with any conclusive educationally significant rise. In fact, the means for grades 8, 10 and 11 in 1978 showed no increase over the means for grades 7, 9 and 10 in 1977.

In the case of the "critical thinking" attitude scale, the grade 11 mean was significantly higher than the means for grade 7, 8 and 10, but the latter three were very close. One hypothesis which could be considered to explain this is that schooling has little effect on the critical thinking - open-mindedness attitude, but that the better "critical thinkers" are the ones who remain at school to pursue the Higher School Certificate.

Further research is needed into the relationship between scores on the attitude scales and other measures. The achievement tests used in this and the 1977 study ranged from general science tests through to specific physics, chemistry and biology tests, which were not separately identified. The correlations between the conservation scale and a specific "knowledge of conservational issues" scale and between the "critical thinking" scale and tests of understanding or knowledge about scientists and their work would also be worth investigation.

REFERENCES

McBURNEY, S. The Junior Secondary Science Syllabus: An investigation; Investigation 4 - attitudes to attitudes Directorate of Planning Services, N.S.W. Education Department, 1978.

GARDNER, P.L. Science curricula and attitudes to science: A review. The Australian Science Teachers Journal, 1975, 21(2), 23-40.

KERLINGER, F.N. & PEDHAZUR, E.J. Multiple regression in behavioural research. New York: Holt, Rinehart and Winston, 1973.

MADDOCK, M.N. Formal schooling and the attitude of Papua New Guineans towards natural phenomena - a replication and extension study. Research in Science Education, 1977, 7, 83-95(a).

MADDOCK, M.N. The attitude of Papua New Guineans towards natural phenomena: A replication and extension study. Report to the Education Department of Papua New Guinea. The University of Newcastle, 1977(b).

MADDOCK, M.N. An exploratory study in school assessment: The relationship between achievement and attitude scales. Research in Science Education, 1978, 8, 127-134.

SECONDARY SCHOOLS BOARD. Syllabus in Science, years 7-10: Statement of aims and content. N.S.W. Department of Education, 1973(a).

SECONDARY SCHOOLS BOARD. Syllabus in Science, years 7-10: Assessment guide, 1973(b).

SULLIVAN, M.J. Student goals, attitudes and achievement. The Australian Science Teachers Journal, 1975, 21(2), 41-44.

THOMSON, P. The affective outcomes of science teaching: Some teacher opinions. The Australian Science Teachers Journal, 1975, 21(2), 5-21.

Research in Science Education, 1979, 9, 75-84

SCIENCE ACHIEVEMENT AND ATTITUDES AT PRIMARY AND SECONDARY INTERFACE

Colin Power

INTRODUCTION

In Australia each summer sees over 220,000 boys and girls of about twelve years of age leave the familiar surroundings of their primary school to enter the new territory of secondary school. That many students do experience difficulty in adjusting to the change has long been recognised. But the nature, extent and origins of the problems encountered by students are not well understood. For instance, little is known of the precise nature of the environmental changes encountered or of the long term effects of such changes on student satisfaction towards school in general or their attitudes towards particular school subjects. In this paper, the relationships between the environmental changes faced by students in transition and the changes in their attitudes towards science are reported.

The theoretical perspective adopted in this study is one which seeks to take both person and situation variability into account in understanding human behaviour. To this interactionist conceptualization of human behaviour has been added a developmental frame of reference forming what Hunt and Sullivan (1974) have called a "developmental inter-actionist" theory. Such a perspective has considerable appeal in studying transition problems. Its essential concerns may be expressed mathematically as an extension of the classic Lewin formulation:

$$\frac{dB}{dt} = \frac{d(P,E)}{dt}$$

This expression emphasizes that as a result of changes in environments (E) and interactions between persons (P) and environments, the persons themselves undergo changes which may cause them to behave (B) differently on subsequent occasions (t). The model also implies that an environment may be growth promoting when it is congruent with a person's needs, and growth inhibiting when there is lack of congruence.

In general, the research which has looked at the transition between primary and secondary school has not attempted to study the dynamic process of interaction between student and school. The focus has been on the stresses and adjustment problems experienced at the point of transfer, rather than the longitudinal changes occurring during the transition period. The majority of students at the point of transfer expect, and in fact encounter, identifiable problems of adjustment (Harris, 1961; Lavenburg, 1963; Murdoch, 1966; Nisbet & Entwistle, 1969; Alfrey, 1966). The most serious, most frequent and the most persistent stresses reported arise out of difficulties in adjusting to the academic environment of high schools (Alfrey, 1966; Lutz, 1966; Winton, 1969). The nature and intensity of the transfer problems experienced by students appear to be a

function of social background, sex, age, ability and personality characteristics (Blackburn, 1966; Nisbet & Entwistle, 1969; Yutzy, 1969).

Neither the short-term transfer nor the long-term transition effects of the actual changes in the environment encountered by students in transition have been systematically studied. Following analytic and individual trait theory, the research has tended to focus on the person, and, in particular, on the presence or absence of psychological symptoms of maladjustment rather than the process of adaptation or the extent of growth. For the individual student, his parents and his teachers, the increasing variability of classroom environments creates new problems of adjustment and new possibilities for optimizing development. In the study, the developmental consequences of the environmental changes encountered by different types of students are explored, these consequences being mapped in terms of changes in student satisfaction, confidence, attitudes and achievement. This paper reports focusses on the effects on attitudes to science of such changes. Those interested in the wider effects are referred to the major report of the project (Power & Cotterell, 1979).

METHOD

The subjects were 516 pupils from 34 Grade 7 classrooms in 12 Brisbane primary schools in 1976 who became Grade 8 pupils in 41 classes in four Brisbane high schools in 1977. A detailed longitudinal study was made of what happened to these pupils. A Student Opinion Scale measuring student satisfaction with aspects of schooling, student confidence and student attitudes towards science and other subjects was administered early (March) and late (November) in last year of primary school and the first year of secondary school. In addition, measures of the learning environment of the classrooms in which these students were located in Grade 7 and Grade 8 were obtained using the Tisher and Power (1975) adaptation of the Learning Environment Inventory (Anderson, 1971) and the Classroom Activities Questionnaire (Steele, House & Kerins, 1971). Fourteen primary and nine secondary classes were observed using the Brophy-Good Dyadic system (1969) and interviews conducted with 134 Grade 7 and 56 Grade 8 pupils and their teachers. Measures were also obtained of student ability, school achievement and personality (the High School Personality Questionnaire, Cattell & Cattell, 1969). Details of the procedure and instruments are given in Power and Cotterell (1979).

Complete data were available for 383 students in transition. Data analyses reported here are based on this data set.

RESULTS: REPEATED MEASURES ANALYSES

The effects of transition on student satisfaction and attitudes towards science were assessed by univariate repeated measures analyses of variance procedures. In the first series of analyses, these procedures were used to determine the effects of:

A: the type of secondary classroom attended (open-plan vs. self-contained)

B: the type of primary classroom attended (open-plan vs. self-contained)

C: sex (male vs. female)

D: occasion (early 7 vs. late 7 vs. early 8 vs. late 8).

The first three factors were regarded as fixed effects, and the last as a nested random factor. The conservative F test (Winer, 1962) was used in testing the significance of effects since this test is designed to avert difficulties associated with the assumption of equal covariances among pairs of observations.

Table 1 and Figure 1 summarize the results of these analyses with general satisfaction with school, confidence in school-related situations and attitudes towards science as the criteria.

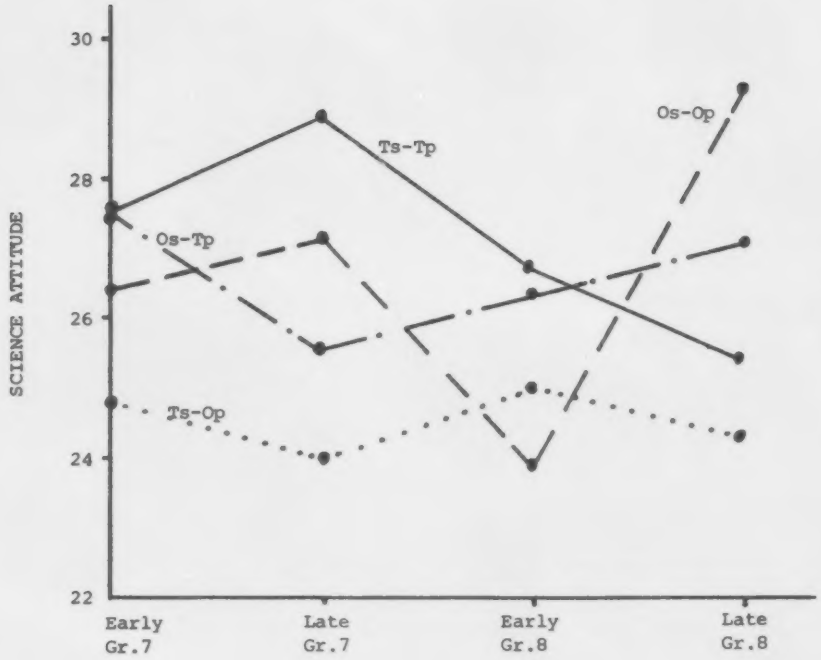
TABLE 1

Repeated Measures ANOVA, Attitudes Towards Science.

Source	df	MS	F	P
Between subjects				
Secondary classroom (A)	1	259.5	2.2	0.13
Elementary classroom (B)	1	512.6	4.4	0.03
Sex (C)	1	1035.4	8.9	0.00
A x B	1	536.1	4.6	0.03
A x C	1	109.7	0.9	0.66
B x C	1	0.3	0.0	0.96
A x B x C	1	16.2	0.1	0.71
(Sub W/I GRPS)	342	116.5		
Within subjects				
Occasion (D)	3	78.9	1.7	0.15
D x A	3	283.8	6.3	0.00
D x B	3	153.7	3.4	0.01
D x C	3	33.4	0.7	0.53
D x A x B	3	198.6	4.4	0.00
D x A x C	3	12.1	0.2	0.85
D x B x C	3	51.1	1.1	0.33
D x A x B x C	3	26.8	0.6	0.62
(D x Sub W/I GRPS)	1,026	45.1		

FIGURE 1

Interaction Effects, Science.

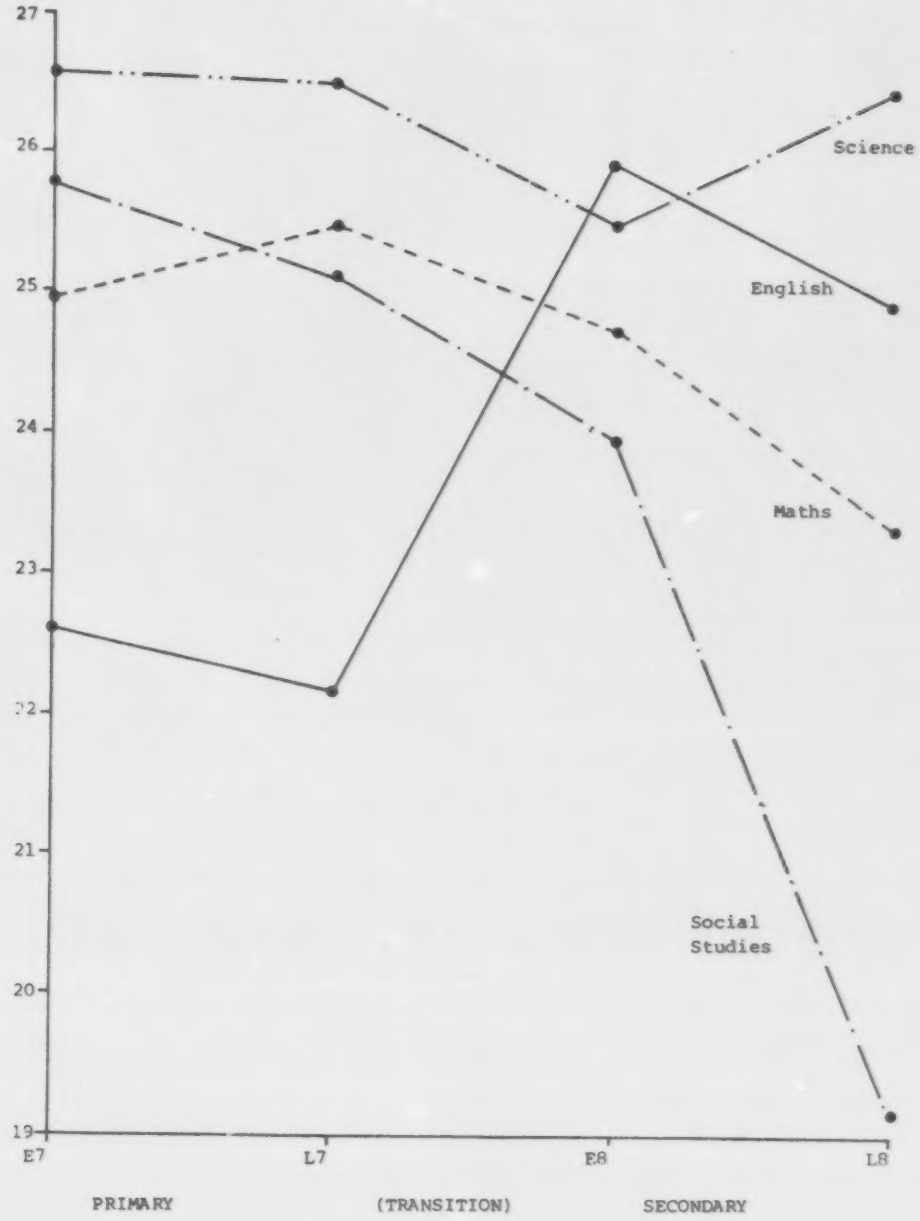


Key: Os = Open plan secondary
Op = Open plan primary
Ts = Self-contained secondary
Tp = Self-contained primary

The results indicate that for the particular classrooms sampled in this study, students were more satisfied when in self-contained primary classrooms, but less satisfied when in self-contained secondary classrooms. During the period there were small initial gains (between 0.2 and 0.4 S.D.) in satisfaction towards school followed by small losses, with satisfaction peaking early in Grade 8. Despite the stresses reported in interviews at the point of transfer, there were moderate gains in student confidence. A significant interaction effect was obtained indicating that whereas the attitude towards science of students in self-contained high school classrooms begins to decline towards the end of the transition period, there is an improvement in attitudes of students in open-plan high school classrooms.

FIGURE 2

Changes in Attitude Towards School Subjects.



It is also of interest to note that, overall, students' attitudes towards science as a subject were consistently higher than their attitudes towards other school subjects as shown in Figure 2. On examining students' ratings on selected items of the semantic differential scale (used to measure attitude towards science and other subjects) it seems clear that at the end of Grade 8, the relatively positive attitude towards science arises largely out of the fact that the subject is seen to be particularly "useful" by students ($\bar{X}$ = 5.61, S = 1.87) as well as "good" ($\bar{X}$ = 5.38, S = 1.85) and "interesting" ($\bar{X}$ = 5.31, S = 1.97). In comparison, maths is also seen as "useful" ($\bar{X}$ = 5.66, S = 1.89) but not as "interesting" (X = 4.08, S = 2.24), while social studies is seen as becoming progressively less "useful" ($\bar{X}$ = 4.39, S = 1.99) and less "interesting" ($\bar{X}$ = 4.59, S = 2.01). Ratings of difficulty and clarity were not markedly different.

That lessons in science tended to be more interesting and the content seen as more useful was confirmed, for the most part, in observations of lessons and interviews. In contrast to mathematics and social studies lessons, students in science classes were more often on task and particularly active in small group, inquiry-oriented activities - largely it would appear as a consequence of the widespread use of A.S.E.P., J.S.S.P. and "Man in Nature" materials in Grade 8 classrooms, and the use of J.S.S.P. and activity cards in Grade 7 classrooms. The activity-oriented science lesson was more prevalent in open-area secondary schools than in the self-contained classrooms of conventional high schools. In the latter, more lessons consisted of conventional teacher-centred approaches, chalk and talk and correction of assigned work.

While the responses of pupils on the semantic differential indicated little difference in attitude towards science between primary and secondary school, the lack of change is probably more apparent than real. Our interviews and observations indicated that, overall, students had little contact with science in primary school and that their positive attitudes were more in an anticipation of what science lessons might be than a reality:

> Science is better than last year. We do many more experiments. I enjoy working in a group and at my own pace. I learn more and it's fun.

> The teacher did everything in science last year. We didn't do much science anyway. Only about once a month. We do everything this year. I like Science and Miss A. is a good teacher.

Whereas some students had doubts initially (cf. dip in attitudes early in Grade 8), in most cases Grade 8 science teachers appear to have convinced the doubting Thomases:

> First term, I disliked science. Now I'm taking an interest and doing my best.

> (Q. What changed?). Well I thought I was interested in history. Last year, I liked it a lot. I thought science was sort of selfish and useless. Now I find it very interesting. Science like English and Maths is needed in every job.

Student attitudes towards science tend to revolve around whether the teacher is seen as "good" or "weak" or "vindictive." "Good" teachers are able to inspire interest, explain things well, are seen as fair, effective managers and with a sense of humour.

> I enjoy science especially Mr. S. He cracks jokes.
>
> Mr. B. is a good teacher. He really helps us with our work and understands us. He's fun.
>
> Science is good for some kids but I don't really like it because Mr. D. picks on me and that upsets me. He doesn't help us much. If we don't know, he gets us to find it out. If we can't, he says it doesn't matter.

REGRESSION ANALYSES

A principal component analysis revealed that the 12 environmental scales used could be clustered into three basic dimensions, according to the extent to which the classroom is perceived as:

> E1: involving i.e. supportive, purposeful and fostering participation
>
> E2: unstructured i.e. disorganised, lacking structure and cohesion
>
> E3: pressured i.e. pressurized, competitive and task-oriented.

Using the means of the 134 classes involved in the study as the unit of analysis in analyses of variance, it was found that primary classrooms were perceived to be more involving and more structured than secondary classrooms ($p < .01$). Contrary to pupil expectations and teacher opinions, there were no differences in academic pressure (difficulty, competitiveness, academic press).

The 9 personality measures also were condensed to four dimensions on the basis of a principal components analysis. These were:

> P1: a self-sufficiency factor i.e. resourcefulness, reflectiveness, aloofness
>
> P2: a maturity factor i.e. emotionally stable, socially confident, self-controlled
>
> P3: an assertiveness factor i.e. aggressive, competitive, attributes success and failure to other factors rather than self
>
> P4: a conscientiousness factor i.e. persistent, sensitive, insecure.

Regression analysis procedures were used to determine the extent to which student attitudes towards science and science achievement could be accounted for in terms of person characteristics (sex, ability, P1 to P4) and the characteristics of the primary and the secondary classroom environment in which the student was located (E1 to E3). The results are

summarized in Table 2.

TABLE 2

Regression Analyses, Attitudes Towards Science and Science Achievement.

Predictor	PRIMARY ATTITUDE SCIENCE R^{2+}	β	r	SECONDARY ATTITUDE R^2	β	r	SECONDARY ACHIEVEMENT SCIENCE R^2	β	r
Person									
Sex	03**	-15	-18*	00	-02	-07	01	08	10
Ability	05*	11	13	07**	24	26**	25**	47	49**
Self-sufficiency	05	06	01	07	08	-01	25	-04	-09
Maturity	06	11	14*	08*	13	14*	25	04	10
Assertiveness	09*	-17	-14*	09	-09	-05	26	-04	-01
Conscientiousness	09	03	00	09	-03	02	26	07	17*
Environment									
P. Involving	09	-03	05	11*	-21	-14*	26	-02	-03
P. Unstructured	10	02	06	11	-07	06	26	-08	00
P. Pressured	16**	-28	-29**	11	-08	-07	26	03	01
S. Involving		n.a.		11	05	09	28	10	12
S. Unstructured		n.a.		11	03	-02	28	-06	-10
S. Pressured		n.a.		13	-06	-11	28	10	03

* $p < .05$ ** $p < .01$

+ Increment significant at specified level

n.a. not applicable

The results indicate that in the primary school, positive attitudes towards science are more likely to be found amongst able, non-assertive boys located in relatively unpressured classroom environments. In Grade 8, the differences between boys and girls are small, the major determinant of both attitude and achievement being ability. Mature students tend to display more positive attitudes towards science, while conscientious students are more likely to achieve. The nature of the environment, as measured by student perceptions, is not strongly associated with outcomes, although the more involving the primary environment from which he or she

came, the less favourable the attitude towards science of the Grade 8 student. It would appear that transitions from highly involving primary school environments to less involving secondary environments are likely to have adverse effects on student attitudes. A number of additional regression analyses were undertaken to check on the possibility of interaction effects as revealed by product terms in the regression equation. The results did not indicate the presence of any such effect.

CONCLUSION

The results in this study provide some evidence indicating that there are no major problems confronting students in transition in the schools sampled so far as science is concerned. The majority of students enjoy science at both levels, and, contrary to popular belief, there is no sharp drop off in student attitude as the year wears on.

The maintenance of student satisfaction and positive attitudes seem to be related to the nature of the curriculum materials used and the quality of the interpersonal relationship between teacher and pupils. It would appear that in the open-area high schools involved in this study, the materials used (largely A.S.E.P.) and the identification of the teacher with the school and the students were more conducive to student satisfaction than those in use in the more conventional schools.

That relatively few of the environmental measures used were found to be associated with student outcomes is a little puzzling. One suspects that in part the reasons may lie in the limitations of the instrument and in part in the lack of variance in student attitudes towards science across the transition period. A careful analysis of what actually happens in science lessons and interviews with the participants reveals that whereas most students are satisfied, the reasons for this satisfaction are as rich and varied as the patterns of adaptation utilized by students in transition.

REFERENCES

ALFREY, E.C. Articulation practices and problems between elementary and junior high schools in Nebraska. Ed. D. Thesis, The University of Nebraska Teachers College, 1966.

ANDERSON, G.J. The assessment of learning environments: a manual for the Learning Environment Inventory and the My Class Inventory. Halifax, Canada: Atlantic Institute of Education, 1971.

BLACKBOURN, J.M. Articulation problems in the junior high school. Ph.D. Thesis, East Texas State University, 1966.

BROPHY, J. & GOOD, T. Teacher-child dyadic interaction. Austin, Texas: Research and Development Centre for Teacher Education, 1969.

CATTELL, R.B. & CATTELL, M.D.L. Handbook for the Junior-Senior High School Personality Questionnaire. Champaign, Ill.: Institute for Personality and Ability Testing, 1969.

HARRIS, S. A study of the relationship between the sixth grade students beliefs about junior high school, his concept of self and others, and his socio-economic status. Unpublished doctoral dissertation, Florida State University, 1961.

LAVENBURG, F.M. An analysis of articulation problems of seventh and tenth grade students with implications for administrative action. Ed. D. Thesis, Rutgers University, 1964.

LUTZ, J. The relative effectiveness of a formal versus an informal orientation program for reading sixth grade students concerns about going to junior high school. Ed., D. Thesis, Temple University, 1967.

MURDOCH, W.F. The effect of transfer on the level of childrens' adjustment to school. Unpublished Thesis, University of Aberdeen. Cited in Nisbet, J.D. & Entwistle, N.J. The Transition to Secondary Education. London: University of London Press, 1966.

NISBET, J.D. & ENTWISTLE, N.J. The transition to secondary education. London: University of London Press, 1969.

POWER, C.N. & COTTERELL, J. The transition between primary and secondary school. Report to E.R.D.C., Flinders University, 1979.

STEELE, J., HOUSE, E. & KERINS, T. An instrument for assessing instructional climate through low-inference student judgment. American Educational Research Journal, 1971, 8, 447-466.

TISHER, R. & POWER, C.N. The effects of classroom activities, pupil perceptions and educational values where self-paced curricula are used. Report of A.A.C.R.D.E. Project, Monash University, 1975.

WINER, B.J. Statistical principles in experimental design. New York: McGraw Hill, 1962.

WINTON, J.L. Articulation problems between elementary and secondary schools affecting Indian and non-Indian students in Cherokee County, Oklahoma. Ed. D. Thesis, University of Arkansas, 1970.

YUTZY, J.A. Articulation problems as perceived by pupils, teachers and administrators in seventh grades of selected junior high schools in the Indiana Public School Study Council. Ed. D. Thesis, Ball State University, 1970.

Research in Science Education,1979,9,85-93

INVESTIGATING STUDENT UNDERSTANDING OF BASIC PHYSICS CONCEPTS USING AN INTERVIEW-ABOUT-INSTANCES TECHNIQUE

Roger J. Osborne and John K. Gilbert

INTRODUCTION

The basic premise of this paper is that words in themselves have no meaning; always the meaning of a word lies in the mind of the hearer or reader. Consequently a very difficult problem in teaching and learning arises from the fact that a student may give to a word a meaning which is subtly different from that intended. For example, if a student has a more limited, extensive or idiosyncratic meaning for a word than the teacher, the full implications of the teacher's message will not be realized. Unfortunatly often neither the student, nor the teacher, is aware of the problem because the student believes he or she has understood. Hence the basic reasons for consequential learning difficulties, hidden in unidentified earlier misinterpretations, may remain quite obscure.

The novice physics student is presented with a particular problem by certain everyday words which are used in a subtly different way in physics, e.g. words like velocity, power, displacement and even words like potential and electric current. While all physics teachers are aware of this problem, lack of time and the constraints of normal classroom interactions prevent a teacher from acquiring a real awareness of the word meanings that students bring with them to physics classes and take with them from these classes.

From a theoretical analysis of current views on concept formation (Markle & Tiemann, 1970; Klausmeier et al., 1974; Gagne, 1977) we have developed a method for exploring students' meanings of those common words which are used in physics in a subtly different way. We have explored the potential and limitations of the method in an exploratory study which involved 50 students, ranging in age from 7 to 48 years, in individual half-hour interviews. This paper outlines the method and the findings of these exploratory investigations.

A METHOD FOR INVESTIGATING STUDENT UNDERSTANDING OF BASIC CONCEPTS

Klausmeier, Gagne, Markle et al. suggest that the key measure of concept attainment is the ability of an individual to categorize properly, instances not previously encountered as instances, or non-instances, of the particular concept. However it appears to us (Osborne & Gilbert, 1979) that to investigate fully a student's concept of a word it is necessary not only to ask a student to classify instances and non-instances of the concept but also to establish the basis upon which the student has made each particular decision.

The above viewpoint has led us to develop the following individual interview method as a practical way of investigating a student's concept of a word. The method involves the use of a set of simple drawings on individual cards to represent the instances and non-instances to be discussed. For each drawing the student is asked "in your meaning of the word.....do you consider this to be an example?" Following this, the reason for his or her decision is discussed. In the case of more abstract concepts, e.g. mass, energy, the student is asked if the depicted instance contains the concept, or "what questions would you want to ask me (the interviewer) before you (the student) could decide whether or not this was an example of the concept?"

In proposing the method we considered that the individual interview situation would allow:

(i) flexibility in discussing reasons, or lack of reasons, for a particular categorization,

(ii) information to be gained about student understanding in situations where the student is unsure how to classify an instance,

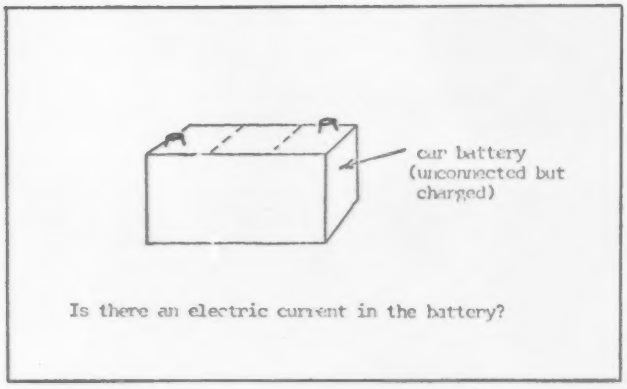
(iii) the student to ask questions to clarify perceived or actual ambiguities before attempting to classify an instance.

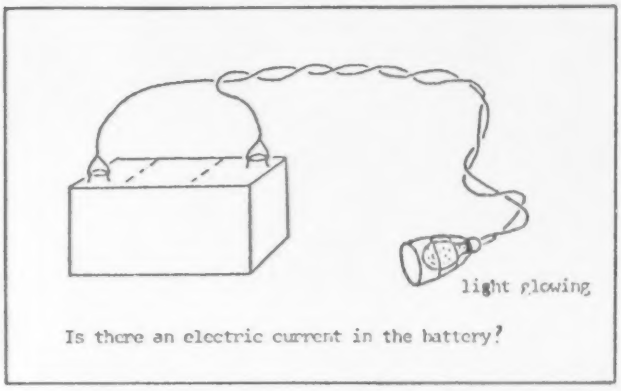
We also realised that the success of the proposed interview-about-instances method depended on the choice of an appropriate set of exemplars and non-exemplars of the concept, so that the critical aspects of students' understandings could be explored. The theoretical structure of the concept, the attributes of the concept, the relationships between the concept and other concepts, various contexts associated with the concept, and comments from experienced teachers about misunderstandings all needed to be considered in the choice of a suitable set of instances. Further the discussion of these instances with students in a pilot study was necessary if an effective set of instances was to be obtained.

Investigating Student Understanding of Electric Current.

In the exploratory investigations student meanings for the words "work" and "electric current" were studied. The purpose of these investigations was to explore the limitations and potential of the method and were not designed to make any generalizations about student meanings from these words. To illustrate the interview-about-instances method, specific aspects of the "electric current" investigations are detailed here.

"Electric current" is an abstract concept without directly perceptible instances. Accordingly illustrations of objects and situations which may or may not contain instances where electric current is present were used as the instances to be discussed in the interview and the student was asked "is there an electric current in...?" or "what questions would you need to have answered before you could decide if there is an electric current in...?"





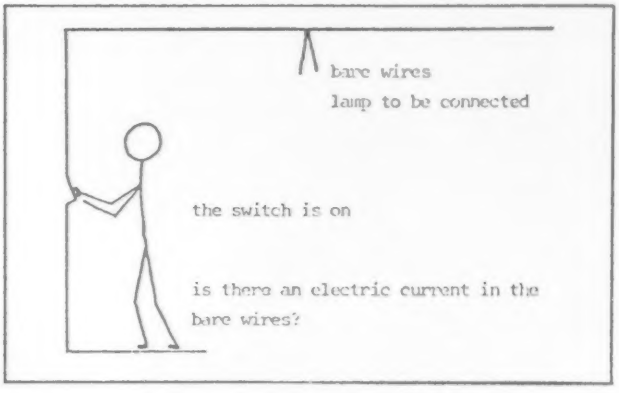


Fig 1: Examples of some of the electric current cards

An initial set of instances was developed from the theoretical perspectives of "attributes" and "contexts":

(i) everyday situations involving none, some or all of the attributes of electric current, e.g. charge but no flow, flow but no charge, and so on.

(ii) types of electric currents as classified in physics textbooks, e.g. elementary particle beams in a vacuum, electron flow in metals, ion flow in gases and liquids, magnetization currents, polarization currents, movement of macro-charges, and so on.

The instances and the order of the instances were modified after an initial discussion of the instances with young children and colleagues.

Over 30 students varying in age from 7 to 48 years were interviewed about electric current and of these 27 were formal 30 minute recorded interviews. Five students were trainee science teachers over the age of 21, thirteen were 16 to 18 year old students studying A-level physics while the nine younger children had little or no formal education in physics or physics related science.

One of the main findings about students' understanding of electric current which we obtained from the trial interviews was that although some 16 to 18 year olds displayed an understanding of electric current similar to those of physicists, in terms of the instances discussed, others showed understandings very similar to those of the 7 to 13 year old children who had received no formal teaching about electric current.

For example in the case of the unconnected battery (first instance in Fig. 1), there were some good answers such as the following from a 17 year old student:

No, I don't think so....there may be some minor thing that was not supposed to be there.

However what might be called the container theory was prevalent at all ages. Eight out of the 22 students aged 7 to 18 years, given formal interviews, made comments like:

Yes, if it has not all been used up. (16 year old)

There is but it is not flowing. (16 year old)

Yes...the same as when it is connected up. (18 year old)

Yes...well the thing is built for electric power. (10 year old)

Yes...because if you connected it to a light the light would glow. (14 year old)

While 16 of the 22 students considered that there was a current in the battery when the light was connected, (second instance in Fig. 1), four A-level physics students had a view represented by the quote from an 18 year old:

> No, there is no electric current in the battery...
>
> there is just an electric current in the wires...
>
> there is just something in there that produces electric current.

Sixteen out of the 22 students stated that there would be an electric current in the bare wires in the third instance depicted in Fig. 1. While 5 out of the 6 who said "no" were A-level physics students, one was an 11 year old girl who stated, "No, it has to have the lamp connected." However 8 out of 13 A-level candidates as well as all but one of the 9 children aged 7 to 14 years stated there was an electric current in the bare wires. Compare the answers of a 17 year old and a 7 year old respectively:

> Yes...because if you put a bulb there it lights up and there is a current. (Q: If the bulb is not there?) Yes...it is still there. (Q: Where is it?) Just in the wires.
>
> Yes...because the light works on electricity and the light has just been cut off. (Q: So it is just there is it?) Yes.

Compare also the answers of a 10 year old, 14 year old, and 18 year old respectively:

> Yes but the electric current would not be doing anything useful...the electric current would flow up to the ends of the wires...but there would not be any electric current coming out...it would flow around in the wires...perhaps back again.
>
> I think there would be electric current in the wires, yes... it is not being put to any use it is just flowing through. (Q: Where to?) I am not sure...it is just flowing through to the ends of the wire...it is not being put to any use... the ends would store up energy.
>
> Yes there is an electric current in the wires, yes... (Q: How much compared to when the lamp was connected up?) I would have thought the current would be the same.

One other finding about students' understanding of electric curent which we gained from the interviews with students was that, as with the concept "work," students tend to have a much more personalized view of physics, and their explanations are related to human experience and feelings, far more frequently than would be expected from a mature physicist. Some examples of this are given in Table 1.

On the other hand explanations which were perhaps more physics-type explanations in that they tended to be in terms of a formal statement relating concepts, often contained glaring misconceptions. As in the case of the "work" interviews, students would apply physics rules and physics knowledge without any real understanding of the physical situation or without even any attempt to relate the situation to their intuition. These students seemed to have learnt to suppress intuition and simply to apply rules in isolation from the real world in which they lived. For example:

TABLE I

Explanations Related to Human Experience or Feelings,
From Three 18 Yr Olds, Two 16 Yr Olds and a 14 Yr Old.

Q: Is there an electric current in the leads to a light bulb socket (bulb missing)?
A: Yes, because I had a shock once when I cleaned the bulb.

Q: Is there an electric current between two electrodes in a liquid if there is a current in the wires from the electrodes?
A: No, I think not because for the simple reason that we have tried the experiments, you know it is called electrolysis, and I have dipped my fingers in electrolytes when the circuit was connected and I didn't feel anything.

Q: Is there an electric current in a strike of lightning?
A: Yes, you will feel a shock - you will be killed if lightning strikes you indicating a large current.

Q: Is there an electric current in a writer's arm as he writes?
A: No, I don't think so - because you don't feel a current.

Q: Is there an electric current from the sun to the earth?
A: No, if there was when it reached people it would earth through the people and they would all die.

Q: Is there an electric current in an electric fence?
A: Yes, because it is called an electric fence and if you put a hand on the wire you would get an electric shock.

TABLE II

Two Replies by A-Level Physics Students to the Question.
"How Would You Define an Electric Current?"

A: Um - I would normally say it is a flow of electrons but it isn't always - because it might be through a solution - then it isn't a flow of electrons - so - um - I suppose a flow of charge from one place of a different electric potential to another place - um. Q: You inferred earlier you need a circuit to have a current? A: Yes - well - I don't know if it has to be in a circuit or not. Q: You said lightning was not a current? A: No - well I suppose - I don't know - I mean I wouldn't have thought of it as a current but from the definition I have given for a current - it would seem to be a current with me. Q: You have never had to express it in words? A: I might have but I would have just said an electric current is a flow of electrons - but obviously that isn't the only thing it is.

A: Um - these (cards) have made me think a lot more about it so - I am not sure about it - what I (had) thought is that if you had a wire which conducted electricity then the electrons in the wire move rapidly along the wire - but (now) I am not sure what it is - I couldn't give a definition.

The car battery has a voltage and V = IR so the battery (unconnected) must have a current. (18 year old)

Lightning has a high voltage and therefore a low current. (18 year old)

In the electric current interviews students were asked, after they had discussed 12 cards depicting instances and non-instances of electric current, to try to give an explicit statement about what they meant by electric current. This request did not provide any further insights into student understanding of electric current except with respect to the relationship between their interpretation of instances and the definition they gave. Four students who displayed a physicists's view of electric current as shown by categorization of instances had great difficulty formalizing a statement (for example see Table II). On the other hand a 16 year old and an 18 year old who gave clear and reasonable definitions, (e.g. "electric current is a flow of electrons") actually had very serious misconceptions (e.g. "electric current is a flow of electrons and in a lamp they change into light and heat" and "light is electromagnetic waves and electric current is a flow of electrons therefore light must be a flow of electrons." It appears to us from this type of evidence that in general, a student's answer to a request for a definition is likely to bear little relationship to a student's understanding as exemplified in the discussion about instances. If one agrees with Klausmeier, Gagne, et al. that the ability to classify is the key component of concept understanding then classifying instances, and providing valid associated reasons, are central to the investigation of a student's concept.

IMPLICATIONS FOR FUTURE RESEARCH

It is not our intention to give a detailed analysis in this paper of the potential and limitations of the interview-about-instances (IAI) method as we have provided this elsewhere (Osborne & Gilbert, 1979). However, in brief, we consider the method has considerable potential for a variety of reasons:

(i) it enables the student's views to be explored without comparison against external criteria,

(ii) it is applicable over a wide age range and it is non-threatening to students,

(iii) it appears to be more pertinent and penetrating than asking for a definition,

(iv) it has advantages over written examinations in that both the student can ask questions about the questions (e.g. if it is considered ambiguous or it is intentionally ambiguous) and the interviewer can query responses or reasons for lack of student response.

On the other hand there are, naturally, some problems to be overcone:

(i) choosing a limited but adequate set of instances so that various

aspects of a student's concept can be exposed,

(ii) choosing a suitable order for the presentation of instances, bearing in mind that particular instances may influence a student's response to later instances.

On balance, however we consider that the method has considerable potential for investigating concept attainment.

Interviews and discussions with teachers college students and teachers have led us to believe our work is of relevance to teacher education. Aspects of the method can be used to:

(i) evoke discussion amongst teachers about their own concepts and improve their own understandings,

(ii) provide information for teachers about likely student understandings, and

(iii) provide a set of illustrations which could be used in formal teaching or classroom discussion.

Our work has also suggested to us a number of topics related to student development of basic concepts which might profitably be investigated using the IAI method. These include:

(i) the prevalence of egocentric views of the world at various ages, the retention of early intuitive ideas despite formal teaching, and the effectiveness of various curricula for changing students' concepts,

(ii) the inhibiting effects on learning of the differences between the general language which students use and the language of teacher and textbook,

(iii) the orientation (and possible attitudes) of students towards words such as "science," "physics," "chemistry," and "biology" particularly at the stage where students make a decision about what subjects they will continue to study.

Consideration of some of these topics has begun at both Waikato and Surrey.

REFERENCES

GAGNE, R.M. The conditions of learning (3rd edition). New York: Holt, Rinehart and Winston, 1977.

KLAUSMEIER, H.J., GHATALA, E.S. & FRAYER, D.A. Conceptual learning and development - a cognitive view. New York: Academic Press, 1974.

MARKLE, S.M. & TIEMANN, P.W. "Behavioural" analysis of "cognitive" content. Educational Technology, 1970, 10(1), 41-45.

OSBORNE, R.J. & GILBERT, J.K. An approach to student understanding of basic concepts in science. Guildford, IET: University of Surrey, 1979.

Research in Science Education,1979,9,95-101

WORKING MEMORY CAPACITY AND LEARNING DIFFICULTIES

N. C. Kellett

INTRODUCTION

Most models of memory propose a system having at least three components - first, a transient store for sensory input information, finally a long term storage component of very large capacity and one or more intermediate components that have storage and processing functions.

There are models that describe the effects long term memory and learning have on each other (Ausubel, 1968; White, 1977). In general, the relevance of such work to the learning of science is obvious. The situation seems to be different when one considers the intermediate memory component(s). Much of the experimental work in this area has considered "nonsense" learning, and perhaps for this reason its relevance to classroom learning has not been appreciated. While there have been general accounts of the way in which the intermediate memory component(s) constrain problem solving strategies (Lindsay & Norman, 1977), its role in the learning of a discipline such as science does not seem to have been considered.

In this paper, we shall give a brief introductory account of an hypothesis that gives a detailed description of the affect that this memory component has on learning. In general terms, the hypothesis proposes that because of the constraints on the capacity of this component, there is an interactive connection between the information content of a task, students' levels of relevant conceptual understanding and their perception of the difficulty of that task. More particularly, where there is a low level of conceptual understanding, students may perform reasonably well and not complain of difficulty in low information tasks; but will complain of difficulty and will show a dramatic drop in performance in high information situations.

This hypothesis was originally proposed to account for the difficulty secondary pupils in Scotland had experienced in one section of their organic chemistry course (Kellett 1978). While we will use examples from this study to illustrate the proposals of the hypothesis, the hypothesis itself would be generally applicable to classroom learning.

The Capacity of Working Memory.

The intermediate memory component has been variously termed "immediate memory," "short ter memory" and/or "working memory." (Feigenbaum, 1977, for example proposes two systems - short term memory and working memory). We shall use the term "working memory" (which seems to emphasize its dual nature) for the intermediate component. In spite of the different models of memory, there is general agreement that working memory has limited capacity both for storing information and for processing information. The

storage capacity has traditionally been measured as the "short term memory capacity" 7 ± 2 chunks (Miller, 1956). A "chunk" of information is what a particular individual recognizes as a unit.

F G Q N S V L T B M

P R O C E S S I N G

C A T F I S H N E T

The preceeding three lines show three rows of 10 letters; in a sense, each row contains the same quantity of information. However english speakers would recognize the first row as 10 single letter units - 10 chunks, while the second would represent just 1 chunk and the third 3 chunks.

The processing capability of working memory is constrained in that processing occurs in a time sharing mode rather than simultaneously. Messaro (1975) cites a frequently observed example: a young child, having read a text aloud may be quite unable to give the meaning of the text because most of the available processing time is taken up with perceiving the words and phrases to be spoken. When processing aimed at eliciting the meaning of the text could begin, earlier essential words have passed from working memory.

This latter example also illustrates the fact that it is difficult to make a clear distinction between a purely storage overload and a purely processing overload. Baddeley and Hitch (1974) have proposed a model of working memory that allows a partial trade off between storage load and processing capability. They have suggested that working memory contains a "store only" component, and a flexible "central processing space" which can either supplement the storage capacity of the "store only" section, or can be used for processing. Their experimental results indicate that the capacity of the store only component is the traditional 7 ± 2 chunks of information, and that when this load is exceeded, there is a significant decrement in the performancae of cognitive activities. They attribute this decrement to the fact that the central processing space must accomodate the information overflow and that the decreased space available for processing must cope with both the increased load of internal servicing of stored information and also the cognitive processing required by the imposed task.

It is the dual constraint on the capacity of working memory that is important here. In one sense, the constraint is the same for all. However, because the unit of information storage is the chunk - a user dependent quantity - the effective capacity of working memory is different for different individuals (for a specific category of information) and will differ for different types of information (for an individual).

Levels of Conceptual Understanding.

Before stating the hypothesis, we should define the term "level of conceptual understanding." We have used this term for two main reasons. First, presumably because much of the work involving concept learning has

considered learning of simple class inclusion concepts, the terms "concept formation" and "concept acquisition" are formally defined (Bruner et al. 1967, Ausubel 1968) in terms of the formation and testing of hypotheses regarding the attributes of a concept and the process of finding predictive defining attributes that distinguish exemplars from non exemplars. We would argue that terms so defined cannot generally be applied to the learning of science concepts as in general students are given formal definitions (a set of predictive defining attributes) of a concept. It is then the learner's task to build up for himself the generic meaning of the concept, partly through meaningful learning of the definition. Secondly, we would argue that for most science concepts it is not appropriate to treat concept learning as dichotomous - i.e. the learner either "has" or "has not" learned a concept. One would not expect HSC students and post graduate students to have indentical concepts of, say, momentum or chemical bonding. It would seem more appropriate to consider concept learning a transitional process in which higher states of a concept differ from earlier, lower states or levels in terms of their intrinsic generic meaning, their degree of abstraction, the degree linking with other concepts etc. This model of concept learning is adequately described by the term "levels of conceptual understanding."

THE HYPOTHESIS

We propose that task information content, level of related conceptual understanding and perceived difficulty are related in the following way:

1. The number of chunks representing a given body of information will depend on the level of relevant conceptual understanding.

2. The greater the number of chunks used at some stage of a task, the greater will be its adjusted difficulty and the poorer the observed performance.

3. In the limit, no useful information may be obtained if the learner attempts to handle all the information simultaneously.

4. When chunk capacity is exceeded, new information may be obtained using a memory conserving strategy i.e. a strategy that allows sequential consideration of small amounts of information.

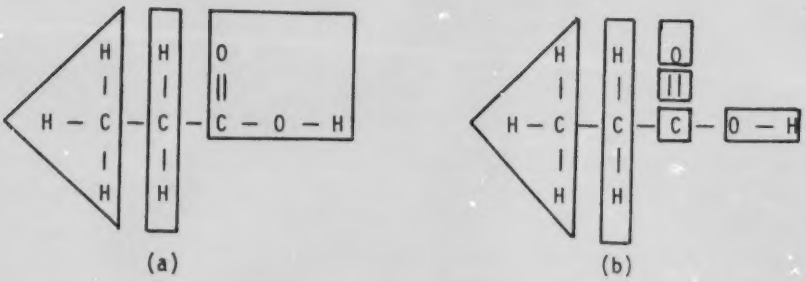
5. Appropriate levels of conceptual understanding lead to an efficient, organized converging strategy. Low levels of conceptual understanding can lead to inefficient, poorly organized strategies, or even arbitrary and diverging strategies.

Chunking and Conceptual Understanding.

Conceptual understanding can affect the number the chunks required first in that it affects the amount of information per chunk and secondly in that it may determine how much information is chunked.

FIGURE 2

Two Ways of Chunking a Formula.



One might reasonably hope that secondary school pupils who had studied organic chemistry for up to three years would consider functional groups to be chemical entities, and hence recognize them as units. This level of conceptual understanding could lead to the chunking indicated in Fig. 2(a). A pupil with lower conceptual understanding could chunk the same formula as depicted in Fig. 2(b). In fact, of a combined sample of fourth, fifth and sixth year Scottish pupils (N = 1292) only 5% recognized the common functional groups as units and demonstrated chunking as in Fig. 2(a). 75% of the sample demonstrated Fig. 2(b) chunking, while 10% of the sample had a single symbol or double bond per chunk. In this case, a higher level of conceptual understanding required fewer chunks to represent a formula because more information was contained in a chunk.

An individual with good conceptual understanding may implicitly declare some of the given information redundant at least in the first instance. For example, in many situations the precise structure of the chain in a carboxytic acid is irrelevant. An experienced chemist could then chunk a formula as (long chain) (acid) or (short chain) (acid); a junior student might feel constrained to chunk all the information given in a formula, hence increasing his memory load considerably.

The nett result of these two effects is that a pupil with a high level of conceptual understanding, by chunking more efficiently and more selectively, is more likely to be able to handle the information content of a task without exceeding working memory capacity than a pupil with a lower level of conceptual understanding.

The Use of Strategies.

It is important to realize that different levels of conceptual understanding lead to different memory conserving strategies, rather than to the use or non-use of such strategies. Pupils with good conceptual understanding may well have an overview of the task as a whole, which enables them to follow a well organized sequence of steps, making a logical selection of the information required at each stage. Pupils with low conceptual understanding may well be quite unable to select appropriate information for step-wise consideration; they may, by default, consider quite arbitrary combinations of information. Where the information content of a task is moderate, learners maybe achieve some success in spite of a low level of conceptual understanding. For example in the organic chemistry study mentioned, students were required to compare pairs of formulae and decide whether or not they belonged to the same family. This is a very simple task for a student who considers functional groups as units. Many pupils who had not achieved such a level of understandiong simply compared the two formulae bit by bit (some spoke of "cancelling" the bits) and decided that the formulae were "different" when they found two bits to be "different." This strategy was inefficient in that they considered irrelevant information and their low conceptual understanding often lead them to a wrong decision. Nevertheless, they had a clear over-view of the strategy they were using, which certainly limited the information should be attended to next. These pupils performed such tasks quickly and did not judge them to be difficult, in spite of their moderate success rate. If the information content of a task greatly exceeds chunk capacity and if the outputs of various steps have to be synthesized, the likelihood of success given a strategy based on low conceptual understanding, becomes very small indeed.

Of course, it may be argued that pupils are generally given rules or solution strategies for difficult tasks (for instance, marked examples of problem types). In a low information situation it is possible to give a simple rule that will cover all possibilities. A pupil with low conceptual understanding may well be able to apply such rules with considerable success. However, the rule required for a high information situation may well exceed chunk capacity for pupils with low conceptual understanding. Thus, the pupils who need to use the rule are those least able to apply it.

Thus, in a low information situation learners may be quite successful whatever their level of conceptual understanding. Higher information situations, however, are sources of learning difficulty for learners with low conceptual understanding.

IMPLICATIONS FOR LEARNING AND TEACHING

Modern syllabuses generally give great importance to understanding principles and concepts. Consequently, many tasks are given to pupils primarily to help them acquire understanding of the concepts involved. The hypothesis outlined above would suggest that the levels of understanding which are the desired outcome of the performance of these tasks are also a necessary prerequisite for the successful handling of the information given. The learner with low conceptual understanding is caught in a vicious circle.

The hypothesis described gives a model which may increase our understanding of the causes of learning difficulties in these educationally important situations. Because of the detailed nature of the relations it proposes, it also suggests guidelines for reducing the effect of low levels of conceptual understanding in handling high information situations. If we wish pupils to increase their level of conceptual understanding in high information situations, we should teach them strategies which enable them to handle the information as though they had conceptual understanding, but which do not themselves depend on the extant level of conceptual understanding. It follows from the relations proposed in the hypothesis that appropriate strategies would:

(a) enable learners to chunk more efficiently;

(b) reduce the amount of information to be considered by a learner; or

(c) provide the learner with simple, efficient strategies for handling information sequentially (these could be solution rules, or strategies that would enable the learner to select salient information more readily).

It would seem that these guidelines may be particularly useful in the sciences and mathematics. In these subjects it is often possible to manipulate the information content of a task without necessarily affecting its educational value. A detailed consideration of the implications of this hypothesis is beyond the scope of this paper. However, we hope that the description of a cause of learning difficulties may help in remedying the problem it identifies.

REFERENCES

AUSUBEL, D.P. Educational psychology; a cognitive view. New York: Holt, Rinehart and Winston, 1968.

BADDELEY, A.D. & HITCH, G. Working memory. In The psychology of learning and motivation: advances in research and theory Vol. VI. New York: Academic Press, 1974.

BRUNER, J.S., GOODNOW, J.J. & AUSTIN, G. A study of thinking. New York: John Wiley and Sons Inc., 1967.

FEIGENBAUM, E.A. Information processing and memory. In D.S. Norman, Models of human memory. New York: Academic Press, 1970.

LINDSAY, P.H. & NORMAN, D.A. Human information processing: An introduction to psychology, 2nd ed. New York: Academic Press, 1977.

KELLETT, N.C. Studies on the perception of organic chemical structures. Ph.D. thesis, University of Glasgow, 1978.

MASSARO, D.W. Experimental psychology and information processing. Chicago: Rand McNally, College Pub. Co., 1975.

MILLER, G.A. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

WHITE, R.T. A model of cognitive processes. Research in Science Education, 1977, 7, 25-32.

Research in Science Education, 1979, 9, 103-109

MISCONCEPTIONS IN PHYSICAL SCIENCE AMONG NON-SCIENCE BACKGROUND STUDENTS: II

R. D. Linke and M. I. Venz

INTRODUCTION

The purpose of this study was to investigate the level of understanding among, and the nature of misconceptions held by, non-science background students concerning certain fundamental principles of physical science, and to develop a series of diagnostic tests to identify some of the more serious and more prevalent misconceptions.

The study was conducted in three stages. In the first stage, a series of interviews was conducted with non-science background students who had enrolled in various science-based courses (including biological sciences and medicine at Flinders University and nursing at Sturt College of Advanced Education). The purpose of these interviews was to identify the major concept areas in which non-science background students appears to have serious misconceptions (see Linke & Deed, 1977). In the second stage, two of these concept areas were selected as examples with which to establish a method for developing and evaluating a more comprehensive range of diagnostic tests (Linke & Venz, 1978). The areas selected were:

1. Structure of matter, changes of state and solubility; and

2. Electricity.

The third stage of the study was essentially an extension of the second to cover new concept areas and, in a limited way, to refine the earlier methods of test construction and analysis. Three additional topics were examined in this stage:

1. Electromagnetic waves: light and heat;

2. Chemical reactions; and

3. Motion and forces.

For each topic a detailed list of basic principles was defined, beginning with the most basic or elementary principles and progressing to more complex ideas usually taught only in the middle and upper secondary grades. These lists were then examined by selected university staff in appropriate physical science disciplines to check on the accuracy of the stated principles and the logic of their suggested interrelationships. It should be emphasised that these principles were not intended to be strictly hierarchical, but rather to provide a network of logically related ideas that could help in identifying systematic errors arising through basic misconceptions.

Following the definition of basic principles a comprehensive set of questions was devised for each topic. Some of these questions were intended simply to test the students' knowledge or recall of a particular principle, while others concluded a statement of the principle and tested its application. In most cases at least two questions were constructed for a particular principle, either to test different applications of the same principle or to test knowledge of the principle and its application in a different context. The purpose of this was to provide a check on whether, for example, an error in a quantitative application question was caused by lack of knowledge of the particular term or principle, misapplication of this principle, or simply a computational error, and on the extent to which an apparent misconception relating to one question was repeated or followed consistently through to other questions. As with the lists of basic principles, the questions for each topic were submitted for comment to a group of staff with relevant expertise, and following subsequent modification were prepared for the major trial.

Three tests were constructed for the trial, one for each topic. In addition, each test was prepared in two forms, one multiple choice and the other free response. The multiple choice form was intended to provide the basis for the final diagnostic test since this would allow computer analysis of interrelated response patterns. However the limited number of responses for each question could at this stage only be determined intuitively by the project investigators or from the interviews with students conducted during Stage I, the former process being inadequate to reflect the likely misconceptions of non-science background students and the latter, while valuable within the scope of the original exercise, relying on too small a sample of students to indicate the range of potential misconceptions. Thus the free response form was included to identify any common misconceptions not covered in the multiple choice test, and if possible to clarify the reasoning used by students in making errors already identified in the multiple choice test.

The trial tests were administered to a total of approximately 500 matriculation biology students at 10 metropolitan high schools in Adelaide. The schools were chosen to include, but not in any stratified way, students of different ages and from different socio-economic, ethnic and educational backgrounds. The reason for choosing biology students was that most of them had not studied either physics or chemistry beyond the intermediate high school level, yet there were also a substantial number concurrently studying physics and/or chemistry at matriculation level. The latter group was intended simply to provide a basis for comparison of results.

Each student was asked to complete two tests, both of which were in the same form (multiple choice or free response), and those who finished first were often given a third, usually in free response form. Thus there were essentially two different samples of students for the multiple choice tests, one completing both parts of the Electromagnetic Waves test (Light and Heat) and the other completing the tests on Chemical Reactions and Motion and Forces. Because of this it was not possible to make any strict comparison of difficulty levels or patterns of response between the two combinations of tests, although there was in any case little overlap in the basic principles being tested.

CHARACTERISTICS OF THE STUDENT SAMPLE

For each test about half the students had studied neither physics nor chemistry beyond grade 10 (intermediate high school) level; about one-quarter of the students had studied either physics or chemistry (or both) to grade 11 level only, and about the same proportion were currently studying physics and/or chemistry at matriculation level. At the lowest physical science level (grade 10 i.e., the non-science background students) there were approximately twice as many females as males, while at the highest level there were almost twice as many males as females, and the overall proportions were almost the same. This pattern occurred in all four tests and reflects simply the preference among female students for humanities and social sciences rather than physical sciences. The implications of this difference in terms of relative performance in the tests will be examined in the final report on the project.

The age distribution of students differed only slightly between the two testing groups, with the majority of students in the age group 16 to 18 years. The only notable difference was the larger number of students (mostly non-science) above 20 years of age in the second testing group (Chemical Reactions, Motion and Forces).

For each test about half the students had studied mathematics only to grade 11 level (very few had abandoned it after grade 10, perhaps because of school policy constraints), and by far the majority of these were non-science background students. Even among the matriculation mathematics students about one-quarter were of non-science background. On the whole, the highest level of mathematics studied was not less than the highest physical science level, and often it was greater. Thus it would seem unlikely that any difficulty with the questions involved in these tests (which, when they were of a quantitative nature, involved relatively simple calculations) could be attributed to a deficiency in formal mathematics background - the more likely deficiency would seem to be a conceptual one, though this by no means denies the possibility of serious deficiencies in mathematical, or at best computational skills.

DISCUSSION OF STUDENT RESULTS

The following discussion is essentially confined to describing particular patterns of response with comments on the possible reasoning underlying these responses. A more general discussion of overall test performance and of the differences between science and non-science background students will be presented in the final report on the project.

Radiation of Light and Heat.

Many students failed to see the similarity between radiated light and heat waves - for example 97% of students recognised that light is transmitted by waves but only 70% also indicated that heat may be transmitted by waves (some confusion probably occurred here with the other mechanisms for transmission of heat); and while 73% of students were able to recognise the approximate velocity of light through space, less than 25% indicated the same velocity for heat. In particular some students confused

the mechanism for transmission of heat through space with the need for "particles" - descriptions of the process included suggestions that heat would travel "as waves of energy with each particle conducting heat to the next particle in air," and by means of "the vibration of molecules to form a chain reaction" or simply "the vibration of molecules of air."

The emission of light and heat from a candle also produced some differences in response. Only 10% of students indicated that both light and heat would be emitted radially in all directions from the candle flame, while a further 35% were correct apart from indicating a general upward movement of heat (confusing immediate emission with subsequent convection). In addition there were 28% who gave the correct response for light but indicated that heat would only be emitted from the top of the flame and travel directly upwards - again the confusion with convection.

In contrast with these suggested differences between light and heat radiation there were many students who did at least associate the two forms of energy transmission. When asked, for example, whether the dark side of the moon (which is colder than the light side), would radiate heat one of the common responses was "there is no light; thus without light there can be no heat energy."

Characteristics of White and Coloured Light.

Although more than half the students recognised that different colours of light have the same velocity but different wavelengths, only 40% also indicated that white light is actually a composite of different colours - many seemed to consider white light just another colour, and ideas on the relationship between the different colours were often quite confused. Some students thought that the colour of light was in some way related to its strength (intensity), temperature, or ability to penetrate substances.

The process of refraction was also poorly understood, and its relationship to wavelength (hence colour) even less. Of the 40% of students who recognised that the degree of refraction is dependent on wavelength, only 18% (including almost none of the non-science students) understood that different colours would be refracted by different amounts when passing through the earth's atmosphere. And of the 68% who recognised that white light is a composite of different colours only 17% knew the conditions under which a spectrum could be formed. As further evidence of the general confusion about the characteristics of light and electromagnetic waves, although a total of 73% of students recognised the correct velocity for light, only 21% also indicated the same velocity for infra-red, ultra-violet and visible colours of light.

Heat and Temperature.

There were two aspects of heat and temperature that caused considerable confusion. The first was the principle that objects in contact tend toward a state of equilibrium, that is, toward equal temperatures. The examples used to test this were those of a hot spoon placed in a cup of cold coffee and a cold spoon placed in a cup of hot coffee. In all there were 59% of students correct on both questions,

although 7% indicated in both questions that the temperature of the spoon would change beyond that of the liquid, and 4% that the temperature of the spoon would not be affected by that of the liquid. The former misconception was explained by some students in the free response test as the result of "very good conducting properties of metal" leading to an excess concentration of heat.

The second aspect was the notion of reciprocal radiation of heat: the principle that all bodies radiate heat irrespective of temperature. Again there were two questions used to test this principle - only 22% of students were correct on both, while 62% (including a higher proportion of science students) indicated that heat would radiate only from hotter bodies to colder ones and that certain cold bodies (such as the dark side of the moon) would radiate no heat energy at all.

Weight and Mass.

Each of these concepts was tested in terms of its basic definition and then in a simple application question involving the quantitative effect on mass and weight of a satellite placed in orbit. About half the students were able to recognise the correct definitions for both quantities, although only about 20% were also able to apply both definitions correctly. There were 12% who defined both quantities correctly but could not apply either definition, and 6% who interchanged the two definitions, although none of these applied the same (incorrect) definition consistently in the application questions. Again the major problem was confusion rather than consistently applied misconception - this was also indicated by the considerable proportion of students who defined mass as volume (11%) and/or weight as density (19%) without applying appropriately either definition.

Acids and Bases.

Students were asked to define the meaning of the terms acid and base and to interpret two given values on the pH scale. Only 23% of students, including less than 10% of non-science students and more than half the science students, indicated the correct definition for both terms. In addition there were 7% who interchanged the definitions, and 4% gave the same definition for both terms. Many of the students defined an acid as a substance which reacts with water to form a base (33%) and/or a base as a substance which reacts with water to form an acid (36%). With respect to the interpretation of pH values (those given were 5 and 10) only 37% were correct on both; 28% interchanged the interpretations and 19% gave the same meaning for both. In general there was little relationship between knowledge of the definitions and the ability to interpret their meaning on the pH scale.

Interpretation of Chemical Equations.

Here again the major problem was inconsistency and confusion. When given a simple equation for the formation of water from hydrogen and oxygen 73% of students were able to indicate the correct number of reactants but only 19% could calculate the total number of molecules involved in the

reaction - on the latter question 15% gave the number of substances involved, 15% the number of atoms and 17% the number of reactant atoms. A similar pattern occurred in the two questions concerned with conservation of mass - although 40% of students were able to balance an equation by providing the missing entity and to answer a simple computational question, there were almost as many (30% - mostly non-science students) who could solve the problem but not balance the equation, and 10% who could balance the equation but not solve the problem. Likewise the notions of conservation of charge and proportionality in chemical reactions were often handled quite inconsistently, and the general confusion about the meaning of chemical symbols and equations was also apparent in the answers to the free response test.

COMMENTS ON THE RESULTS AND IMPLICATIONS FOR FUTURE DEVELOPMENTS

The overwhelming implication from these results is the same as that heralded in the earlier stage of the project: that although there is a small proportion of students who seem fixed on certain basic and identifiable misconceptions the larger group by far are just confused and erratic in their responses. In some cases they seem to know the terms but cannot apply them; in other cases they can apply a principle without being able to define it. In retrospect this is an obvious finding - but only in retrospect. The assumption underlying the project was that students entering tertiary studies without having studied physical science at senior secondary level would nevertheless have been exposed to a considerable amount of "popular science" through newspapers, magazines, radio and television, and might have developed from this quite firmly held views about certain basic ideas in science. While there can be no doubt of the influence of popular science and of the existence of some fascinating - both reasonable and inexplicable - ideas among these students, the extent of confidence held in these ideas (or the ability to apply them consistently in different and unfamiliar situations) seems relatively low.

This probably means that the intuitive approach used in teaching such students about science through tertiary level "introductory" or "bridging" courses, which is essentially to assume they know little or nothing other than to recognise a few terms, may well be a reasonable one, at least for the majority of students. Indeed, from informal comments made by staff and students involved in such courses at Flinders University this is a reasonable approach. A more serious teaching problem is how to illustrate very basic ideas to otherwise fairly sophisticated students without offending their intellect. This is a social as well as psychological and educational problem, and there is undoubtedly no easy solution.

It seems appropriate in concluding this project to reflect on one of the fundamental definitions and to ask: what exactly is a non-science student? To the extent that it is possible to derive an empirical answer to this question it seems to rest on a comparison of response profiles (across all the sets of related questions) between grade 10, grade 11 and grade 12 science students. From a preliminary analysis of such profiles it seems that the grade 11 science students are more like those with grade 10 than grade 12 science, suggesting that students without grade 12 physics or chemistry are unlikely to have much greater understanding of basic principles than those with no specific study in these disciplines at all

(though they may hold at a rather superficial level a considerably larger scientific vocabulary). Thus although terms may be learned fairly quickly, the development of understanding is a considerably slower process - an important principle to be considered in introductory courses and to lead us to more realistic expectations of the learning capacity of non-science background students.

REFERENCES

LINKE, R.D. & DEED, B.M. Understanding and misconceptions in physical science among tertiary non-science background students: Stage I. Educational Research Unit, Flinders University, 1977.

LINKE, R.D. & VENZ, M.I. Misconceptions in physical science among non-science background students. Research in Science Education, 1978, 8, 183-194.

Research in Science Education, 1979, 9, 111-118

THE IMPACT OF THE INCLUSION OF ASEP MATERIALS ON SOME COGNITIVE OUTCOMES IN DIFFERENT TYPES OF TASMANIAN SCHOOLS

Darrell L. Fisher

INTRODUCTION

At the beginning of 1974 curriculum materials developed by the Australian Science Education Project were available for use in Tasmanian secondary schools. The ASEP materials consisted of units to be used in science classes from grade 7 to grade 10. From its inception one policy adhered to by the Australian Science Education Project was that it would be unwise for any school to use solely ASEP materials. Data gathered during the study indicated that this situation was the policy adopted by Tasmanian schools. The science curricula of many of these schools contain a mixture of ASEP and other materials.

The impact of the inclusion of ASEP materials into the science curriculum on cognitive outcomes was examined, and this paper describes part of this study. In the study the variables investigated were sex, socio-economic status, school type, and the use of ASEP materials. Information on cognitive outcomes was gathered using seven scales from the Test of Understanding Science (TOUS) and the Test of Enquiry Skills (TOES).

This paper discusses the relationship between the use of ASEP materials and the various school types existing in Tasmania.

METHOD

There are three types of secondary schools in Tasmania. Schools which have students in grades 7 to 10 only are termed high schools, while those with students in grades from kindergarten to grade 10 are called district schools. Both schools are administered by the Tasmanian Education Department. District schools are situated in the country areas of Tasmania and usually there is a high school situated in a neighbouring town. Many of the children of higher ability leave the district school at the conclusion of grade 6 to attend the neighbouring high school. Thus there is a depletion of numbers in the district school and consequently a depletion in teaching staff and variety of curriculum. Behrens (1977) has produced data that shows that the children of these district schools are a deprived group. The third type of school belongs to the independent system. These schools are usually church schools with their own councils and functioning independently of the state education authority.

A pre and post test situation was used in which students with ASEP materials included in their science curriculum were compared with those using traditional methods. The major hypothesis being investigated was that if the ASEP materials are really effective then the curriculum in which they are included should be significantly better than the other curriculum on the outcomes being measured. A second hypothesis being

investigated was that there is no difference between the students of the three school types on the outcomes being measured.

At the beginning of the school year a letter was sent to every school in Tasmania that contained secondary classes. Thirty three indicated their willingness to take part (a response rate of 56.9%).

Grade 7 was selected because it is the first year of the secondary education system, and since ASEP materials have been designed for children of grades 7 to 10, students would be exposed initially to ASEP in their grade 7 year.

For the purposes of the study an ASEP class was one in which ASEP materials were included for about 40% of the years' science curriculum in grade 7. In fact the percentage of time spent on ASEP materials varied from 20% to 60% but control classes had no experience of ASEP materials at all. When further categorized into school groups the numbers in the sample are shown in Table I.

TABLE I

SAMPLE CATEGORIZED INTO SCHOOL TYPES

SCHOOL TYPE	ASEP	NON ASEP
HIGH	1529	330
DISTRICT	52	102
INDEPENDENT	147	213

Two instruments were selected. The Test of Understanding Science (TOUS), originally devised by Cooley and Klopfer (1963), and modified in Australia by Fraser (1978); and part of the Test of Enquiry Skills, Fraser (1979). There are three scales on the TOUS test named "Philosophical." "Historic-Social" and "Normality of Scientists," each of which measures some aspect of understanding the nature of science. The four scales chosen from the TOES were Library Usage, Scales, Charts and Tables, and Comprehension of Science Reading.

The choice of this set of scales can be justified. Extensive literature analysis, in particular by Klopfer (1971) and Fraser (1974), has shown that understanding the nature of science and enquiry skills are considered by science educators as very important aims for science programmes. Morever each scale has been validated and satisfactory reliability coefficients have been reported (Fraser 1978, 1979).

ANALYSES AND RESULTS

A multiple regression analysis was performed separately for each of the seven learning outcome post tests using a full regression model containing the following 20 terms; the corresponding pre test, the six coded variables required to specify the set of four main effects (curriculum, sex, school type, and socio-economic status) and the 13 coded variables needed to specify the set of six two-way interactions. These variables were entered into the regression equations in three distinct stages. First the variance attributed to the corresponding pretest, secondly the variance attributed to the block of main effects, thirdly the variance attributed to the set of two way interactions. Thus the main effects were investigated first and any interaction effects were estimated in terms of the variance they accounted for over and above that due to main effects. This method is recommended by Cohen (1975). The relevant section of the analysis is summarized in Table II.

Table II shows that the amount of variance in scores on each learning outcome post test accounted for by the full 20 term model ranged from 19.2% to 18.2% and a significant relationship ($p < .001$) existed in each case. The variance in post test performance attributed to the pre test ranged from 16.7% to 25%. The variance due to the block of main effects, beyond that attributed to the pre test ranged from 1.5% and 3.0% and was significant for all seven criteria. The increment in post test variance associated with the set of two way interactions, over and above that attributed to the pre test and the block of main effects, ranged from 0.3% to 1.2%. Results for the set of interactions were significant ($p < .01$ for only three outcomes, namely the "Philosophical Scale," "Charts and Tables" and "Comprehension of Science Reading").

The variance of the block of main effects was partitioned further into the unique contributions attributable to each of the main effects. Table II shows the amounts of variance associated with the main effect school type beyond that attributed to pre test and the other three main effects. The interpretation of the three level variable, school type, involved re-running the relevant multiple regression analysis three times, omitting the subsample of students belonging to a different one of the three levels on successive runs. The full details of the analysis are provided elsewhere (Fisher, 1978).

When pre test and the other three main effects are controlled, the following interpretations emerged for the seven significant findings of the variable school type: independent school students significantly outperformed district school students on all seven outcomes; independent school students performed significantly better than high school students on the "Philosophical Scale" and "Library Usage" but significantly worse on "Comprehension of Science Reading"; and high school students significantly outperformed district school students on all outcomes except the "Philosophical Scale."

Three cases of significant two way interactions emerged each involving the variable school type. The significant sex by school type interaction for the "Philosophical scale" was interpreted by plotting, separately for boys and girls, the mean for each of the three school types, after adjustment for pre test and the other main effects.

TABLE II

Percentage of Variance in Post Tests Accounted for by Pre Test, Main Effects, School Type, and Interactions.

R^2 (%) / SCALE	FULL 20-TERM MODEL	PRE	BLOCK OF MAIN EFFECTS	BLOCK OF TWO-WAY INTER-ACTIONS	SIGNIFICANT INDIVIDUAL MAIN EFFECT SCHOOL TYPE	SIGNIFICANT INDIVIDUAL INTERACTIONS
PHILOSOPHICAL	19.2**	16.7**	1.5**	1.0*	0.3* (I>H,D)	0.4* (sex x school type)
HISTORIC-SOCIAL	26.4**	24.0**	2.0**	0.5	0.8** (I,H>D)	
NORMALITY OF SCIENTISTS	23.7**	21.6**	1.8**	0.3	0.3* (I,H>D)	
LIBRARY USAGE	28.2**	25.0**	2.8**	0.4	1.1** (I>H>D)	
SCALES	27.0**	24.6**	1.7**	0.6	1.0** (I,H>D)	
CHARTS AND TABLES	24.9**	21.0**	3.0**	1.0*	1.3* (I,H>D)	0.5** (curric x school type)
COMPREHENSION	23.0**	19.2**	2.5**	1.2**	1.5** (H>I>D)	0.9** (curric x school type)
MEAN R^2 (%)	24.6	21.7	2.2	0.7		
DEGREES OF FREEDOM	20	1	6	13		

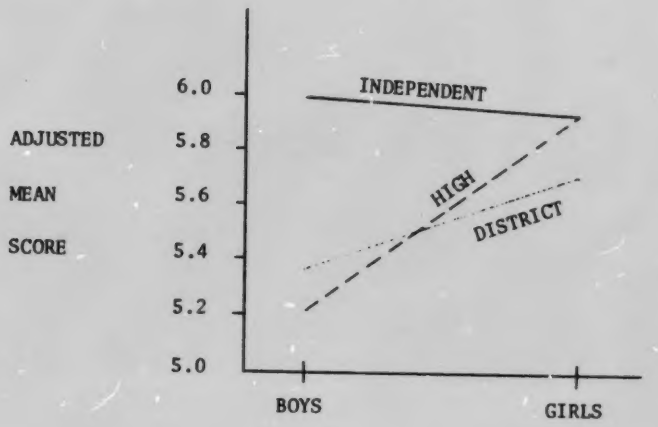
*p < .01

**p < .001

H = High, I = Independent, D = District

FIGURE I

Sex X School Type Interaction: Philosophical Scale.



Performance on the "Philosophical Scale" was almost independent of sex in independent and district schools, while girls outperformed boys on this outcome in high schools.

The second significant two-way interaction occurred between curriculum and school type.

The results showed that high school students performed independently of curriculum type on this scale. In the independent schools the non ASEP group outperformed the ASEP group while the reverse occurred for the district school students with the ASEP group performing better than the non ASEP group.

The third significant interaction occurred in the "Comprehension of Science Reading" scale. Here the interaction was again between school type and curriculum used.

The students with ASEP materials included in the curriculum were comparable in all schools, however when ASEP materials are not included then the district schools performance is inferior to the other two school types. If the two scales "Comprehension of Science Reading" and "Charts and Tables" are considered together it appears that district school students benefit if ASEP materials are included in the curriculum.

CONCLUSION

FIGURE II

Curriculum X School Type Interaction: Charts and Tables Scale.

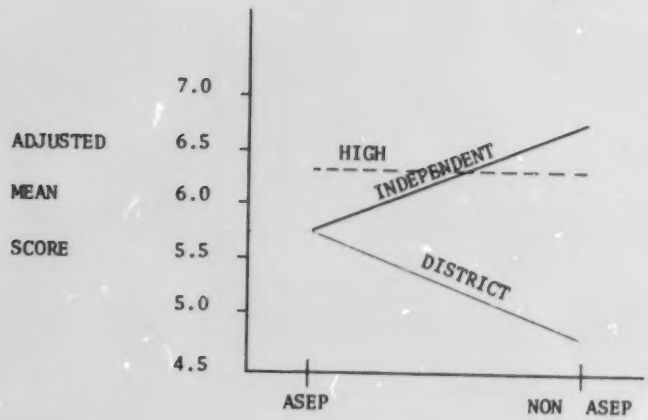
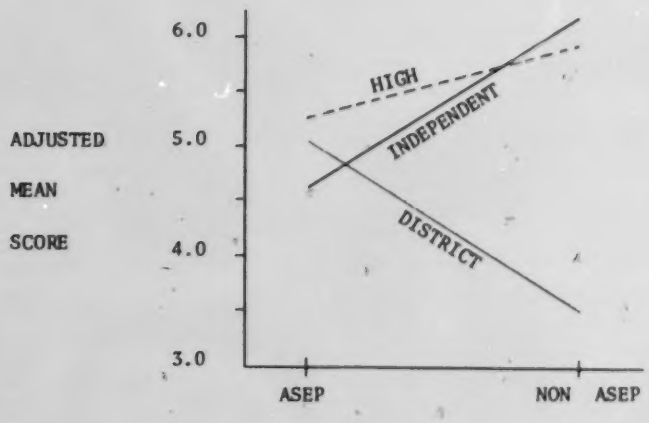


FIGURE III

Curriculum X School Type Interaction: Comprehension Scale.



The most favourable aspect of the inclusion of ASEP materials into the school science curriculum of grade 7 classes in Tasmanian schools emerged from the analysis of two-way interactions. It appears that the students of district schools benefit from the use of ASEP materials. The inclusion of these materials has had no impact on high school students' cognitive achievement. The inclusion of ASEP materials in independent schools had a deleterious effect on the students' learning.

The results for district schools are consistent with the work of Behrens (1977), who conducted a study on country children in Tasmanian schools. District schools are located in the country areas of Tasmania so his findings are relevant. He discovered that country children were a deprived group, noted that they have a lower level of participation in schooling at upper levels, and that their job opportunities are limited. He suggests that in order to reduce the existing inequalities between city and country communities "there is an important need to increase school and work opportunities for country children" (Behrens, 1977, 17). This study suggests that one way in which the children of country areas could be helped is in the provision of science curriculum materials in district schools. The study suggests that ASEP materials are beneficial but of course further research on the learning of science in district schools is desirable. It could be that the schools were so deprived that the provision of any curriculum materials is beneficial and aids the teachers in these schools. There is a need for continued investigation into the provision of science education for children in the district schools of Tasmania.

REFERENCES

BEHRENS, N. Patterns of schooling and work in Tasmania. Panorama, 1979, 1, 14-20.

COHEN, J. & COHEN, P. Applied multiple regression correlation analysis for the behavioural sciences. New York: John Wiley and Sons, 1975.

COOLEY, W.W. & KLOPFER, L.E. The evaluation of specific educational innovations. Journal of Research in Science Teaching, 1963, 1, 73-80.

FISHER, D.L. The impact of the inclusion of ASEP materials on some cognitive outcomes in Tasmanian schools. Unpublished M.Ed. thesis, University of Tasmania, 1978.

FRASER, B.J. Selecting evaluation instruments. In R.D. Linke & L.H.T. West (eds.), Science Education: Research, 1974, 99-112.

FRASER, B.J. Developing subscales for a measure of student understanding of science. Journal of Research in Science Teaching, 1978, 15, 79-84.

FRASER, B.J. Test of enquiry skills. Melbourne: Australian Council for Educational Research, 1979.

KIM, Jae-On & KOHOUT, F.J. Analysis of variance and covariance: Subprograms Anova and Oneway. In Nie, N.H. et al. (eds.), SPSS Statistical package for the social sciences, 2nd ed. New York: McGraw-Hill, 1975.

KLOPFER, L.E. Evaluation of learning in science. In Bloom, B.S., et al. (eds.), Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill, 1971.

LABOVITZ, S. Criteria for selecting a significant level: A note on the sacredness of .05. American Sociologist, 1968, 3, 220-222.

Research in Science Education, 1979, 9, 119-121

ASPECTS OF CHANGE ASSOCIATED WITH THE CONCEPT OF EVOLUTION IN THE AUSTRALIAN SCIENCE EDUCATION PROJECT

Rod Fawns

Four different but overlapping approaches were used to investigate the following related questions:

1. How are the various aspects of the concept of evolution developed in the ASEP units.

2. What formal language appropriate to these aspects is developed in the ASEP units?

3. Could science teachers adapt and extend ASEP units to develop various aspects of biological change?

4. Do current junior science syllabi in Victoria use certain ASEP units in emphasizing the evolutionary synthesis?

No simple answers were expected. The construction of the argument is analogous to Darwin's in the sense that it rests on the "logic of possibilities" described by Hull (1966-7).

The first approach to the content analysis was to "checklist" the content and level of 15 of the most likely ASEP units against the Dedman (1976) sequence of development in each of the five aspects of evolution.

As for Carss and Clarke (1972) the second approach to the problem of describing development in the content analysis was the simplistic one of examining the frequency of occurrence and distribution in the materials of certain index words. A young teacher was employed to look carefully at the printed word in the same ASEP units. She was asked to use Roget's Thesaurus to find synonyms for the index words used formally in the teaching of the concept of evolution and to look for these also. Each unit which was referenced for these words and a comment written to explain the sense in which each word was used in each instance.

A third approach to structural content analysis was made which worked with conceptual plans of each unit in order to discern its purposes and emphases. The guide to ASEP, Teachers Guides and the published bibliography for each unit were important here. A flow chart or concept map, including extensions, for each unit were constructed by a teacher. Here he sought to represent the major goals for each unit expressed in the various guides to the Bibliography relating the suggested goals in such a way as to display the potential value of the unit for teachers interested in developing one or other aspects of biological change. Although a static portrayal, the concept map materializes facts as commodity and permits a single proposition to be made of relationships and non-relationships amongst these facts or assertions (McDonald-Ross, 1972).

In the fourth approach to content analysis over 150 Victorian science teachers (mainly co-ordinators) representing as many schools were invited as part of an in-service programme to portray in another "concept map" what they thought were the major conceptual emphases in their junior science programmes. They were asked to describe each of the units, topics or themes they used at each year level by name and to show by drawing inter-connections between these where they felt core emphases were developed over the four years.

A Few Propositions for Further Study.

On the evidence considered collectively, certain difficulties confront the teacher interested in developing through ASEP units an understanding of biological change.

1. The examples offered of adaption and change are referred to only in passing and their brief treatment for example of adaption seem to reinforce rather than challenge the misconception Dedman found to be very common, that adaptation is an attribute of the organism. Many units (being directed to an ecological unity) are attempting many things at once.

2. The examples of intraspecific variation and selection are almost all human, and controlled. They describe social more than physical features. The impression gained is that the Lamarkian paradigm would most likely be reinforced rather than challenged in these units (Neubergen, 1966).

3. In the Level 1 units diversity, at least, is presented in many forms and more complex patterns described in Stage 2. Notions of adaptation, environmental change, selection and continuous change seem not to have been introduced in any way which would allow them to be used to explain the diversity. Paradoxically the problem may result from the ASEP commitment to an ecological synthesis in these biology units. Developed in this way organisms and relationships seem to be stable rather than changing. The cyclic changes that are represented also refer to states albeit dynamic, rather than transformations. It is a static view of the environment.

4. Two of the three units which bring the formal language together - 312 "Where Did Humans Come From" and 302 "The Australian Scene" are amongst the least purchased of all ASEP units.

The other unit 304 contains much formal language of genetics. It emphasizes variations in humans and each person's uniqueness. It explains the mechanism of uniqueness and described herd and plant improvement programmes. The notion of natural selection and evolution as phenomena would have to be extensions of this unit - they are not central to its purpose.

5. "Evolution" was mentioned in less than one in four "syllabi maps" and in those appeared as an isolated topic or as an aspect of genetics. In the landscape of science in most Victorian schools "change" is nowhere near as prominent a feature as "energy" and

"interactions" (Vallance, 1977).

REFERENCES

CARSS, B.W. & CLARKE, J.A. A content analysis of the core of an ASEP unit: "Pushes and pulls" examined. Australian Science Teachers' Journal. 1972, 18(4), 59-61.

DEDMAN, J.A. The structure and development of concepts associated with the topic of evolution in secondary school boys. Unpublished Ph.D. thesis. Chelsea College, London University, 1976.

HULL, D.L. The metaphysics of evolution. British Journal for the History of Science, 1966-7, 3, 309-337.

McDONALD-ROSS, M. Behavioural objectives and the structure of knowledge. APLET Conference, Institute of Educational Technology, The Open University, Bletchley, Bucks., 1972.

NEUBERGER, H.T. Conceptualization of change held by ten and eleven year old children. Journal of Research in Science Teaching, 1966, 4(3), 180-181.

VALLANCE, E. The landscape of "The great plains experience": An application of curriculum criticism. Curriculum Inquiry, 1977, 7(2).

Research in Science Education, 1979, 9, 123-132

SCIENTIFIC LANGUAGE AND THE TASMANIAN HIGH SCHOOL CHILD. THE TENDENCY TO GENERALISE

P. P. Lynch

INTRODUCTION

In our studies of the ability of high school pupils to recognize concept definitions (Lynch et al., 1978), one feature that emerged from an examination of the pupils' free writing was the extent to which concept words were expressed at different levels of meaning. For instance, the pupils wrote that a solid "was like ice," "had a fixed shape," "was a form of matter." Many of the written forms were recognizable as being fairly standard interpretations, common to many text books and typical of classroom dialogue on the subject.

The recognition or even the production of written descriptions does not necessarily imply understanding but word recognition must surely be at least a pre-requisite to understanding. For instance, it is difficult to see how pupils could understand a written passage like that presented below if they cannot recognize the relevant concept words for what they are.

> As the charge density is very great at a sharp point there is a very strong electric field near it. In air there are always a few electrons present which can be accelerated by the field to such a speed that they collide with air molecules; other electrons are knocked out and positive ions are produced. The released electrons continue the ionising process. (Jardine, 1972).

The implications of levels of meaning for adolescent learners have been discussed by Peel (1975) with reference to sentence preference tests which were non-subject specific. We decided that it would be interesting to reopen this area of investigation, to consider the levels of meaning required by science texts, and examine to what extent pupils showed a preference for thinking at the different levels.

After careful scrutiny of the children's writing and high school texts in science we decided to limit our investigation to three recognizable levels of meaning, illustrated with reference to the concept word "solid" as follows:

> 1. Membership: in which the concept is understood in terms of a single or group of easily recognized object(s), e.g. ice is a "solid."
>
> 2. Partial Association: in which the concept is defined in terms of a specific characteristic or attribute, e.g. a "solid" has a fixed shape.
>
> 3. Generalisation: in which the concept is defined in terms of a generalisation, involving a framework of reference considered to be

particularly appropriate, e.g. a "solid" describes a form of matter.

In the case of the example for membership it will be appreciated that the word ice could be considered to cover both "identifiable instances of ice" and the notion of "iceness," though the teaching emphasis in high school science would almost certainly be on the former. More difficulty is experienced in defending the category partial association which could be considered to be a more limited "generalisation," particularly in the case of the example chosen. However the majority of partial associations in the test deal with symbolic representations, quantifications or aspects of practical measurement.

In practice, dialogue and written materials use all three levels but the highest level, that of generalisation would seem to be the level of meaning necessary for a full understanding of most school certificate texts. That is, classifications, correlations and the interdependence of most important concept terms require reasoning at the level of generalisation. There is a level beyond generalisation, termed "abstration" by Peel (1975). At the level of abstraction the concept is defined as "a notion" or "an idea" which serves to rationalise appearances. The concept solid, might be defined as "a scientific notion which..." A concept defined in such terms is at a higher level of meaning than generalisation but the latter would appear to be more fundamental to the understanding of high school science. This presumes, of course, that we are prepared to accept the validity of the written content of standard texts written at this level. In such texts, arguments are rarely presented at the level of abstraction.

After reflection on these particular points we decided to investigate to what extent Tasmanian high school pupils tend to prefer one of these levels of meaning and how these tendencies change with grade.

METHOD

Construction of the Test.

The sixteen concept words, used previously in our study of "recognition of concept definitions" and associated with the theme the nature of matter, were defined as generalisations. The concept word and its defined generalisation are given in Table 1. Some of the original definitions were modified and where this has taken place they are marked with an asterisk (*) in Table 1.

Each concept word was then defined in membership terms, which presented no particular difficulty (e.g. steam is a gas, a sheet of paper has area, sodium chloride is a compound). The concept words "atom, molecule, element, proton and ion" were related to hydrogen or components of hydrogen, at the level of membership. Each concept word was then defined as two different partial associations. In the majority of cases these were linked to terminology, symbolism or measurement procedures.

TABLE 1

The Concept Word and its Defined Generalisation.

Concept Word	Defined at the Level of Generalisation (G)
1. Atom	An atom is a small particle of matter.
2. Gas *	A gas describes a state of matter.
3. Mass	The amount of substance in a body is called its mass.
4. Neutron *	A neutron is a particle found in the nucleus of atoms.
5. Molecule	The smallest particle of a substance that can exist on its own is called a molecule.
6. Volume	A three dimensional space is called a volume.
7. Proton *	A proton is a particle found in the nucleus of atoms.
8. Solid *	A solid describes a state of matter.
9. Compound	Different elements combine chemically to produce a compound.
10. Length	The distance measured along a line is called a length.
11. Element(s)	Elements are substances composed of only one type of atom.
12. Liquid *	A liquid describes a state of matter.
13. Area	An amount of surface is called an area.
14. Mixture *	A mixture contains more than one element or compound.
15. Electron *	An electron is a sub-atomic particle.
16. Ion	An ion is produced when an atom loses or gains an electron.

Note: Items marked (*) indicate a change from the recognition of concept definition test. The changes for electron, proton and neutron are minor; gas, solid and liquid have been related to states of matter; mixture receives a new definition.

The first three items of the resulting sixteen item test are given below, as they appear in the test.

1. Atom
   A. H represents an atom of the element hydrogen.
   B. An atom is a small particle of matter.
   C. About 100 different kinds of atom exist.
   D. An atom is a neutral particle.

2. Gas
   A. A gas is usually less dense than a solid or liquid.
   B. A gas describes a state of matter.
   C. A gas has no surface.
   D. Steam is a gas.

3. Mass
   A. A mass is measured in units such as the gram or the ounce.
   B. The amount of substance in a body is called its mass.
   C. The mass of an object is measured using a balance.
   D. An object such as a 20 cent piece has a mass.

Using the coding G = generalisation, P = partial association, M = membership, the statements associated with the three items are categorised as follows: MGPP, PGPM, PGPM. The items were subjected to group validation by the appropriate matriculation examiners, six experienced teachers and six new teachers. As a consequence of this procedure and a pre-test at a school not included in the sample, a number of items were re-phrased.

The first page of the questionnaire provided an example item (in this case the word "island" was used). The test supervisor informed the pupils that all the sentences in the test items were correct and that they should read carefully through all four sentences and "pick out the one that you prefer as having the most meaning for you when you think about the underlined word." Pupils were then asked to indicate their choice by placing a tick next to the item. The instructional procedures were also written on the first page of the questionnaire.

Subjects and Related Variables.

A sample of six Tasmanian high schools was chosen as being representative of the high school population in Tasmania. It consisted of boys (969), girls (666). Data associated with I.Q. socio-economic status, age, gender and grade were collected at the same time. Full details of the sample and the validity and reliability of the scales of I.Q. and

TABLE 2

Preference for M, G and P by Grades for Each Item.

	Change in M with GRADE	Change in G with GRADE	Change in P with GRADE	Major Preference GRADE 7		GRADE 10	
atom	↓	↑	↓	G	→	G	
gas	-	↑	↑	P	→	P	**
mass	-	↑	-	G	→	G	
proton	-	↑	↑	G	→	G	
molecule	↓	-	↑	G	→	G	
volume	-	↑	↑	P	→	G	*
neutron	-	↑	↑	P	→	G/P	*
solid	↓	↑	↑	P	→	P	**
compound	-	↑	↑	G	→	G	
length	-	↑	-	P	→	G	*
element	-	↑	-	G	→	G	
liquid	-	↑	↑	P	→	P	**
area	↓	↑	-	G	→	G	
mixture	↓	↑	↑	G	→	G	
electron	-	↑	↑	P	→	P/G	**
ion	-	↑	↑	G	→	G	
Total number of items with direction of change	5↓	15↑	11 ↑ 1 ↓				

* = major preference P at Grade 7

** = major preference P at Grade 7 and Grade 10

socio-economic status are discussed elsewhere (Lynch, 1979).

The total test scores (maximum = 16) were computed for individual categories (membership, partial association and generalisation). These scores were as follows: membership (mean = 0.60, S.D. = 1.0); partial association (mean = 4.80, S.D. = 3.0); generalisation (mean = 3.5, S.D. = 1.8).

Application of the Test.

The sentence preference test was given to all the pupils who, one month previously, had completed the recognition of concept definition test. Those who answered the later test correctly were treated as a single group. The data presented, unless otherwise specified, refers to this particular group, though there are circumstances where the responses for all pupils merits consideration.

Before examing the evidence for systematic trends in preference for thinking styles, the reliabilities, and the correlation with I.Q. were calculated.

Reliability was calculated for each response type by application of the Kuder Richardson formula, K-R20:

Membership Scores r = 0.0,

Partial Association Scores r = 0.68,

Generalisation Scores r = 0.65.

Item Analysis.

There are four possible responses (M, P, P and G) for each item, two of which are categorised under the same heading. The percentage response for each of the categories, M, P and G, was calculated for each of the sixteen items. The average percentage response for M, P and G for the total test was also computed.

RESULTS AND DISCUSSION

Individual Items.

The M, P and G percentages at grades 7,8, 9 and 10 were calculated for each preference item. We then examined how the categories changed with grade for each item and what the major response was at both grade 7 and grade 10. These results are presented in Table 2. There are clearly considerable consistencies in both the trends examined.

The preference for membership, where there is a change, shows a decrease with grade. Three items "gas," "liquid" and "solid" receive more than 10% as a membership score. "Gas," surprisingly, is at the 20% level

at grade 10. Both generalisation and partial association are increasingly preferred for each item with change in grade. The only exception is for the concept, atom.

The major response categories at grade 7 and grade 10 shows "gas," "liquid" and "solid" are preferred at the level of partial association even by grade 10. "Electron," "neutron," "length" and "volume" are involved in a preference transition (P - G) over this period of time.

It should be noted that a large precentage of pupils, who had previously recognised a concept definition at the level of generalisation, prefer it as partial association or even in terms of membership. Cognitive preference is not the same as verbal recognition.

Absolute and Relative Trends of Total Percentages.

The absolute trends with grade (and age) are consistently in the same direction for M, P and G preferences. M decreases while P and G decreases. The absolute percentages (including the correct recognition of concept definition) are given in Table 3.

TABLE 3

The Average % Absolute and % Relative Preferences for M, G and P by Grade (Figure in Brackets is the Relative %).

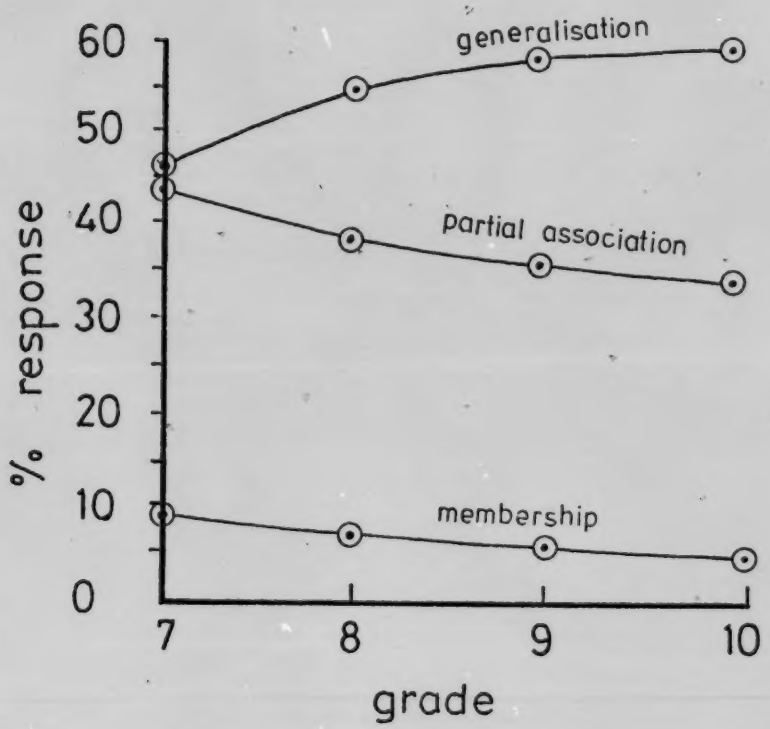
	Grade 7	Grade 8	Grade 9	Grade 10
% M	3.6 (9.2)	2.9 (6.8)	3.4 (6.2)	3.6 (5.7)
% G	18.3 (46.9)	23.1 (54.5)	31.4 (57.0)	37.6 (59.7)
% P	17.1 (43.8)	16.4 (38.7)	20.3 (36.8)	21.8 (34.6)
% Correct recognition of concept definitions (total)	38.9	40.6	51.9	64.6

If the relative percentages of the M, P and G preferences are calculated for each grade then the trend (see Table 3) is somewhat different. Membership decreases more markedly, generalisation increases

but partial association now decreases. A graph of the relative trends (Figure 1) suggests that the M, P and G percentages approach limits, at grade 10, of approximately 5%, 30%, 65% respectively.

FIGURE 1

Relative % of P, M, and G Responses by Grade.



The shape of the curves are interesting and extrapolation suggests that partial association should be the major preference in the last year of primary school and presumably membership should catch up with partial association in the early primary school. These tests were administered towards the end of the grade 7, consequently, grade 7 is seen to be a particularly important transitionary period for high school pupils with regards to preferential thinking.

Correlation with I.Q. Estimate.

Our previous findings for the concept recognition test show I.Q. to be easily the best predictor of performance. The relative correlations of I.Q. with M, P and C categories are of interest not only in their own right but also because they help validate the sentence preference test. High correlations are not to be expected since there are considerable differences between tests of I.Q. (Otis type) and language usage with regard to form and content of the test. The correlations are low but in the expected direction. The correlation coefficients for n = 548 are as follows: M pref. (-0.16); C pref. (0.36); P pref. (0.18).

At this stage we are concerned mainly with establishing the phenomenon of preferential thinking style as it applies to the pool of science words in question. The consistency of the trends for individual items; the inter-relationship of the M, C and P total scores with grade; the correlations with I.Q. estimate, and the reliability of the test would seem to be evidence for the existence of such a phenomenon.

Various colleagues and co-workers have commented that individual items (e.g., solid, liquid and gas) are not surprisingly preferred at the level of partial association in view of the definitions chosen, e.g.:

a gas has no surface

a gas is usually less dense than a solid or liquid.

Such definitions usefully distinguish between solid, liquid and gas even if they are not expressed at the level of generalisation. However, even in the case of solid, liquid and gas the C scores are increasing with grade faster than the corresponding P scores. Indeed, some very intelligent pupils prefer the category membership but the aggregate preferences are definitely in the direction indicated by Fig. 1.

CONCLUSION

Two points associated with the general trends are particularly noteworthy:

1. The marked transition in grade 7.

2. The ceiling in generalisation that is reached by the end of adolescence.

If we take the theme, the nature of matter, as representative then grade 7 (age 12-13), represents a nexus of change in thinking where teaching method and content are particularly important. There is also a maximum performance (implying limits) which we intend to explore further.

REFERENCES

JARDINE, J. Physics is fun. Book 4. London: Heinemann, 1972.

LIFE SCIENCE LIBRARY. Matter. Nederland, N.V.: Time-Life International, 1972.

LYNCH, P.P. The language of science and the Tasmanian high school child. Australian Science Teachers Journal, 1978, 24(3), 78-85.

LYNCH, P.P. Scientific language and the high school pupil. Journal of Research in Science Teaching, 1979 (in press).

McCORMACK, T.H. Chemistry to sixteen. London: University Tutorial Press, 1976.

PEEL, E.A. Predilection for generalising and abstracting. British Journal of Educational Psychology, 1975, 45, 177-188.

Research in Science Education,1979,9,133-143

THE EFFECTIVENESS OF THE USE OF GAMES AS A REVISION TECHNIQUE IN JUNIOR HIGH SCHOOL SCIENCE

Gaye Rosen

INTRODUCTION

This study resulted from questions raised by the author's previous use of games with junior science classes. These included the extent to which the games motivated the students to learn and whether the students learned the concepts that the games were claimed to teach them.

Upon investigating some of the related literature it became apparent that there were questions being raised, and many claims being made, for games. The claims about the beneficial nature of games can be grouped into eight categories:

(i) improved motivation and interest in the subject (Gordon, 1970),

(ii) better cognitive learning of factual information (Abt, 1968); the acquisition of concrete meaning for abstract terms (Gordon, 1970); and better understanding of the processes simulated by specific games (Greenblat, 1973); and improved critical thinking skills (Ellington & Percival, 1977),

(iii) transfer of learning to other situations (Reid, 1976),

(iv) the development of favourable attitudes to social issues such as the conservation of the environment (Heitzmann, 1974),

(v) the development of more positive attitudes to themselves (Greenblat, 1973),

(vi) the development of communication and social skills (Garvey, 1971),

(vii) improvements in the classroom climate and inter-relationships (Heitzmann, 1974),

(viii) benefits specifically for teachers such as increased flexibility (Heitzmann, 1974) and greater sense of their own worth (Tisher, Power & Endean, 1972).

While some of the writers had raised questions about the actual degree of change in motivation and how well the students learned the material, there were many other questions asked that were pertinent to the use of games in the classroom (Hounshell & Trollinger, 1977; Wentworth & Lewis, 1973).

If games are to be an acceptable instructional technique then researchers must supply answers to the following questions:

> What types of knowledge, skills and attitudes can best be learnt using games?
>
> What type of student will do best in a game learning situation?
>
> Are there special organizational details and teacher-personality traits that are more suited for the successful use of classroom games? (Cruikshank, 1976).

No single study could possibly attempt to answer all these questions. However, it was felt that the amount of initial material learnt by students could be measured by an immediate post-test and that cognitive retention could be assessed by using a delayed test. Also there appeared to be some relevant characteristics in students that could be measured in the classroom and which might relate to the successful acceptance of games by the students. It was decided that games could best be used in science as revision aids. Thus there were three basic questions this study tried to answer:

> Do games, used as revision aids, improve the amount of factual knowledge gained initially by students compared with traditional revision methods?
>
> Do games, used as revision aids, improve cognitive retention of factual knowledge compared with traditional revision methods?
>
> Do some of the personality variables present in these students and which might be expected to influence their participation in and learning from games, have a significant effect upon the learning outcome?

The previous research in this area has been fairly limited despite the numerous beneficial claims made for games. Of the studies that have been carried out, virtually all of them have been in the fields of social sciences. Some of the studies have lacked proper controls, had students not selected at random, and have often relied heavily on anecdotal evidence. In addition, many of the studies were inconclusive. A brief summary of the research is given in Table 1.

It is apparent from the table that there is no clear cut evidence that games are a more effective teaching technique and further research is certainly necessary, particularly with respect to science teaching. For semantic clarity, "games" in this study refers to academic games, i.e. games primarily for, or based on, learning as defined by Cruikshank and Mager (1976). This study is different to previous studies involving the use of games because in this study, games were used as revision aids rather than being the sole type of instruction used in teaching a unit where all learning was supposedly achieved via the game situation.

TABLE 1

Results of Studies Comparing Games with Traditional Teaching Methods.

Criterion	Relative Performance of Games Group				Total No. of Studies
	Superior	No Diff.	Inconsistent	Poorer	
Initial Learning	14	9	2	1	26
Cognitive Retention - Recall	4	1	-	-	5
Attitude Change to Issues	7	2	-	-	9
Attitude to Course/ Student Interest	2	2	-	-	4
Decision Making/ Critical Thinking	1	3	-	-	4

Some of the results obtained from research in this area have been criticised on four major grounds:

The apparent interest generated by games might be due to the "Newness" of games rather than the games themselves.

Adequate control groups and any accounting for teacher variables were often lacking.

Little consideration had been given to individual differences in the acceptance of games by students.

There had been little reliability data supplied for the testing instruments used in the studies.

There seems to be justification in the criticism that the introduction of a game into an established teaching programme could be motivating simply because it is a change from the classroom routine. As an attempt to overcome this, Year 7 students were selected for the study. They had been at high school for only 6 weeks when the study began.

METHOD

The study covered two groups, A and B, studying two topics - Digestion and Circulation. Each group had one full class. The experimental design allowed for for Group A being the initial experimental group while Group B acted as the control group for the first topic. For the next topic Group B was the experimental group while Group A acted as the control group.

It is most unlikely that any one teaching method can be all things to all children and it therefore seemed important to look at areas of possible difference among students that may be important in the successful playing of games. The following factors were considered:

(i) need for achievement - as the games used involved competition with reward for winning, it seemed reasonable to hypothesize that students with a high need for achievement would do better in a competitive situation than students with a lower need for achievement.

(ii) need for affiliation - as the games used in the study involved groups of students playing together it seemed reasonable to hypothesize that students with a high need for affiliation might find in games a better learning situation than students who had little need for affiliation and who might therefore prefer a more solitary learning situation.

(iii) cognitive levels in terms of Piaget's theory.

If games, as claimed earlier, do give concrete meanings to abstract terms and concepts then games covering abstract processes, e.g. circulation, should be of great assistance to concrete thinkers but not so much to formal thinking students who are already capable of handling some abstract concepts.

The test used to measure both need for achievement and need for affiliation was the Self Selection Scale developed by Boyd (1976). This scale, on a test-retest analysis with a one-week interval, gave the following reliabilities:

(a) 0.76 for need for achievement for boys, 0.74 for girls;

(b) 0.83 for need for affiliation for boys, 0.70 for girls.

The test used to assess cognitive levels in science was the "Understanding in Science Test" developed by Tisher and Dale (1975). This test when compared to the clinical interview procedure designed by Piaget and Inhelder gave a "77% agreement and indicates that the test is a useful alternative to the clinical interview."

Each of the end-of-topic tests was presented to a panel of three experienced science teachers to assess the validity of the content and its suitability for the age of the students involved. After minor changes in the wording, the tests were administered. The test entitled "Digestion" had a reliability of 0.68 (Kuder Richardson 20), and "Circulation," a reliability of 0.64.

Several additional factors were taken into consideration to try to ensure the internal validity of the study. One factor taken into account was the "history effect," i.e. were all test subjects experiencing the same, or at least very similar, circumstances throughout the period of the study. Secondly, the "experimental mortality," i.e. how were students who missed some of the test accounted for in the study. Thirdly, the similarity of the experimental groups in terms of the type and method of games played. The importance of these types of factors are discussed by

Tuckman (1978).

(i) History effect - this has two aspects: (a) similarity of learning experiences: to try to ensure this a very detailed programme for study was given to the teachers. The classes covered the same objectives and carried out the same practical work over the same time intervals; (b) carry-over effects from the first game: to try to minimize any after effects, (e.g. possible increase in motivation due to the use of games), a "buffer zone" involving a three week unit on "Skeleton & Joints" was used between the two parts of the study, with normal revision methods and only one test given.

(ii) Experimental mortality: if a study missed either an immediate or delayed test for a given topic he/she was eliminated for that topic.

The games used were both commercial games developed by Longmans (1975). The games were on biological topics and were board games in which players moved around the board by the number obtained on a spinner.

It should now be apparent that the collection of the data was carried out so that in this study there would be a maximum internal validity. Therefore this study should hopefully avoid some of the criticisms that have been levelled at other studies involving the use of games.

Specific Hypothesis to be Tested.

Hypothesis 1: Games used as revision aids improve the initial learning of biology content relative to traditional revision methods.

Hypothesis 2: Games used as revision aids improve cognitive retention of biology content relative to traditional revision methods.

Hypothesis 3: Students with a higher need for achievement are advantaged in a game learning situation as compared with traditional learning situations.

Hypothesis 4: Students with a higher need for affiliation will do better in a game learning situation as compared to traditional learning situations.

Hypothesis 5: Students who are concrete thinkers benefit more than formal thinking students in a game learning situation as compared with traditional techniques.

Data Collected for Each Subject.

The data collected on each student included:

(i) sex - as a categorical variable; 1 = male, 2 = female.

(ii) Piagetian or cognitive level - as a categorical variable; 1 = early concrete, 2 = concrete, 3 = early formal, 4 = late formal.

(iii) need for achievement (referred to as nAch) as a continuous variable from 0 - 18.

(iv) need for affiliation (referred to as nAff) as a continuous variable from 0 - 18.

(v) scores on immediate and delayed tests for Digestion.

(vi) scores on immediate and delayed tests for Circulation.

Analysis Used.

Multiple regression was used to generate an analysis of variance using the "Experimental Design Approach" as discussed by Kerlinger and Pedhazur (1973), and McArthur, Fraser and West (1975). This design is analogous to a five-way analysis of variance in which the 5 independent variables are group, cognitive level, sex, need for achievement and need for affiliation. The multiple regression approach however, has several advantages over the traditional analysis of variance. Firstly, it permits the use of the continuous independent variables of need for achievement and need for affiliation without the loss of data that would result if the variables were categorized as would have been necessary if the traditional five-way analysis of variance was carried out. Secondly, it allows a least squares solution to the disproportionate cell data that exists (Fraser, 1975).

For each analysis, the full regression equation contained nine predictor variables. These included the five main effects of group, cognitive levels, sex, need for affiliation and need for achievement. The other four predictor variables came from the interactions formed between groups and each of the other four independent variables. These four interactions were generated because the research hypotheses 3 to 5 required such interactions to be considered. There were no formalized expectations for the other possible interactions from the main effects and therefore they were not generated.

In each analysis, the dependent variable was the individual student's test score. In terms of hypotheses 1 and 2 the immediate test results and delayed test results are essentially independent of each other. Therefore four analysis of variance tables were generated and the results of these are discussed below.

RESULTS

As the second part of the study was a replica of the first part, but with the groups interchanged, it was thought best to bring the results of the two parts of the study together. This allowed consideration of results of the two immediate tests separately from the two delayed tests and therefore made comparisons between both parts of the study considerably easier. Table 2 shows the results of the analysis of variances for both sets of immediate test results.

TABLE 2

ANOVA Results for the Immediate Tests

Variables	Digestion			Circulation		
	df	MS	F	df	MS	F
Group	1	84.877	9.26**	1	6.168	1.49
Cognitive level	1	11.179	1.22	1	16.295	3.94
Sex	1	0.946	0.10	1	11.932	2.89
nAch [A]	1	10.114	1.10	1	0.162	0.04
nAff [B]	1	17.567	1.92	1	0.054	0.01
Group x nAch	1	24.369	2.66	1	8.080	1.96
Group x nAff	1	0.059	0.00	1	0.431	0.104
Group x Level	1	2.721	0.30	1	7.407	1.79
Group x Sex	1	3.904	0.43	1	0.943	0.29
Error	46	9.168		46	4.135	

** $p. < 01$

To simplify the interpretation of the results, it was decided to divide each of the continuous variables into two categories about the mean. For cognitive level, levels of concrete thinking and of formal thinking were collapsed into two categories: concrete or formal.

Hypothesis 1: Games used as revision aids improve the initial learning of biology content relative to traditional methods.

This hypothesis was supported in the first part of the study. The experimental group actually scored higher marks in both parts of the study, but only for Digestion was this difference significant at the 5% level. This tends to support the usefulness of games as revision aids. This result is similar to those results seen in Table 1.

As stated earlier, the results of the two delayed tests were also considered jointly. Table 3 shows the results of the analysis of variances for both sets of test results.

TABLE 3

ANOVA Results for Delayed Tests

Variables	Digestion			Circulation		
	df	MS	F	df	MS	F
Group	1	10.215	2.44	1	16.580	3.33
Cognitive level	1	1.444	0.34	1	7.697	1.55
Set	1	6.340	1.51	1	4.167	0.84
nAch	1	10.990	2.62	1	0.810	0.16
nAff	1	0.000	0.00	1	0.029	0.00
Group x nAch	1	4.826	1.15	1	0.608	0.12
Group x nAff	1	6.411	1.53	1	5.700	1.14
Group x cognitive level	1	25.291	6.03*	1	12.298	2.47
Group x sex	1	26.841	6.40*	1	10.388	2.09
Error	46			46	4.982	

Hypothesis 2: Games used as revision aids improve cognitive retention of biology content relative to traditional revision methods.

This hypothesis was partly supported in the Circulation segment of the study where the games group had a mean that was higher than the control group and this approached significance ($p<0.1$). There were no significant differences in the means for Digestion. This apparent lack of support for the hypothesis is against the trend previously found, where, of the 5 studies reported, four found that the games groups had retained significantly more knowledge than the control groups. However, as all these studies used games for the full length of the instruction period as opposed to this study, where the games were used strictly as a revision technique, it may indicate that a longer time period for game playing is necessary to achieve the desired effects. Based on the results of this

study, the situation is still equivocal with respect to this hypothesis.

> Hypothesis 3: Students with a higher need for achievement are advantaged in a game learning situation as compared with traditional learning situations.

The graphs in Appendix 1 suggested that the higher need for achievement students were advantaged in a game situation as compared to the control group on both immediate and delayed tests, however, the differences were not significant. The fact that the same trend appeared in all four test situations may indicate that future research using larger numbers may show significant interactions. However, the results of this study indicate that the effectiveness of games is independent of a student's need for achievement.

> Hypothesis 4: Students with a higher need for affiliation will do better in the game situation as compared to traditional learning situations.

There were no significant interactions or consistent trends in the results of the study. This would seem to indicate that the effectiveness of games is independent of a student's need for affiliation.

> Hypothesis 5: Students who are concrete thinkers benefit more than formal thinking students in a game revision situation as compared with traditional techniques.

There was one significant interaction with respect to this hypothesis and this was for the delayed test on Digestion where the formal thinking students were advantaged by the games. This was the opposite result to that expected. Further it was also the opposite to that seen in the initial test on Digestion, where the concrete thinking students appeared to be advantaged. For Circulation, the trend in both tests was for the formal thinkers to gain benefit from the game situation. It would seem that further research is needed in this area before the situation can really be understood. The evidence at this point in time indicates an equivocal position on this hypothesis.

There was no formal expectation of either sex being advantaged in the game situation because the games had no inbuilt bias for either sex. There were no significant interactions in either of the immediate tests although there seemed to be a trend that indicated that the girls were being advantaged in the game situation. In a delayed tests, there was a significant interaction on Digestion showing that the girls were more advantaged than the boys in the game situation. For Ciculation, the trend was the opposite of this with the boys apparently being advantaged by the game, although this trend was not significant. These results would seem to indicate the possibility of sex bias and further research may be warranted.

The teachers involved in the study made some anecdotal reports on the students' acceptance and interest in the games. They reported that later in the year transport strikes, zone athletic carnivals and similar events, often reduced the number of students in a class to such a level as to prevent continuation of the present topic. In this case the students, when offered a free choice of doing any science activity they wanted, usually

asked to be allowed to play one or both of the games. They repeatedly played the games, often as many as four times in a double period without discipline problems arising. It is doubtful if students would have shown the same enthusiasm and self control for any other revision techniques. No matter how good a colour filmstrip or television film might have been, it would be very unrealistic to expect students to want to view the programme repeatedly.

The fact that students chose to voluntarily play scientifically oriented games gives some encouragement for their usage in the classroom.

CONCLUSIONS

It should be pointed out that the results of this study showed the students did just as well, and sometimes better, in a gaming situation when compared with traditional revision situations. Therefore games were as good as, and sometimes a better revision aid than, other more traditional means such as video tapes and slides. Hence games could be useful as a means of providing variety and active student involvement in the classroom. Further, the acceptance of games did not appear to be dependent, to any large extent, on personality factors such as need for achievement, need for affiliation, and cognitive level. This seemed to be an indication that games were suitable for use with a broad spectrum of students.

REFERENCES

ABT, C. Simulation games in learning. In Boocock, S. & Schild, E. (eds.). California: Sage, 1968.

BEARD, R. Teaching and learning in higher education. Harmondsworth: Penguin Books, 1970.

BLOOMER, J. What have simulation and gaming got to do with programmed learning and educational technology. Programmed Learning and Educational Technology, 1974, 10(4), 224-232.

BOYD, R. Measuring need achievement and need affiliation in the classroom. Australian Psychologist, 1976, 11, 185-192.

CRUIKSHANK, D.R. Simulation: New direction in teacher preparation. Phi Delta Kappa, 1976, 48, 23-24.

CRUIKSHANK, D.E. & MAGER, R. Toward theory building in the field of instructional games and simulations. Programmed Learning and Educational Technology, 1976, 13, 5-9.

ELDER, C.D. Problems in the structure and use of education simulation. Sociology of Education, 1973, 46, 335-354.

ELLINGTON, H. & PERCIVAL, F. Educating 'through' science using multi-disciplinary simulation games: Programmed Learning and Educational Technology, 1977, 14, 117-126.

FRASER, B.J. The impact of ASEP on pupil learning and classroon climate. Research in Science Education, 1975, 5, 1-12.

GARVEY, D.M. Educational aspects of simulation. In Tansey, P.S. (ed.). London: McGraw Hill, 1971.

GREENBLAT, C.S. Teaching with simulation games. Teaching Sociology, 1973, 62-83.

HEITZMANN, W.R. Educational games and simulations. Washington: National Education Association, 1974.

HOUNSHELL, P. & TROLLINGER, I. Games for the classroom. Washington: National Science Teachers Association, 1977.

KERLINGER, P. & PEDHAZUR, E. Multiple regression in behavioural research. New York: Holt, Rinehart and Winston, 1973.

McARTHUR, J., FRASER, B. & WEST, L. Some issues in the simulation of ANOVA by multiple regression techniques. Research in Science Education, 1975, 5, 199-207.

McLEOD, J. Gaming...is it simulation? Simulation, 1971, 228-229.

REID, N. Simulation, games and case studies: Education in Chemistry, 1976, 82-83.

TANSEY, P.J. Educational aspects of simulation. London: McGraw Hill, 1971.

TISHER, R. & DALE, L.G. Understanding in science test. Hawthorne, A.C.E.R., 1975.

TISHER, R.P., POWER, C.N. & ENDEAN, L. Fundamental issues in science education. Sydney: John Wiley and Sons, 1972.

TUCKMAN, B. Conducting educational research. New York: Harcourt, Brace Jovanovich, 1978.

WENTWORTH, D. & LEWIS, D. A review of research on instructional games and simulations in social studies education: Social Education, 1973, 37, 432-440.

Research in Science Education, 1979, 9, 145-147

AUDIO-TUTORIAL APPROACH—A MULTI-MEDIA OR AN INDIVIDUALIZED METHOD OF INSTRUCTION?

Abdul G. Khan

INTRODUCTION

A source of concern to any teacher is the wide range in student learning abilities and rates of learning. There are numerous methods or patterns of adapting education to individual student needs and differences (Edling, 1971; Norris et al., 1975). Individualized study can be varied from extreme structure (that requiring little teacher direction) to extreme flexibility (that requiring little teacher direction). One approach for individualization is the audio-tutorial (A-T) system as developed at Purdue University by Professor Postlethwait (Postlethwait et al., 1969). This system uses an audio-tape to guide pupils through integrated learning sequences, which include laboratory experiments, printed study guides, visual materials, and demonstrations. The author (Khan, 1979) has shown that the cognitive achievements of student-teachers at Goulburn College of Advanced Education in Introductory Biology can be significantly increased by utilizing the A-T instructional approach.

In the present study an attempt was made to find out at least some trends in the attitude of the students towards the audio-tutorial approach after they completed their course. Of the five techniques Bloom et al. (1971) listed to evaluate affective objectives, the rating-scale closed-item anonymous questionnaire technique was employed in the present study.

The responses reported here are based on the views of 232 student teachers who have completed their introductory science courses using the A-T approach, during the years 1976, 1977 and 1978.

RESULTS

In summary, the following opinions were evident:

Approximately two thirds of the student teachers surveyed (60%) rated the A-T approach a good teaching/learning method. The rest of the respondents regarded it a bad method (25%) or were indifferent (15%).

Comparison between attitude towards the conventional lessons and the A-T sessions showed that 62.5% of the respondents found the latter more effective. The rest were indifferent.

There was some difference of opinion between the respondents as to the enhancement of self-reliance or confidence by the A-T method. Some (37%) regarded that the method has increased their self reliance, others (50%) were indifferent and still others (13%) responded strongly negative.

In response to the attitude towards the small assembly sessions in the A-T method, the majority (62.5%) showed positive attitude. They favoured the sessions with regard to inter-personal communication and to integrate theory with practical work. The rest of the respondents were indifferent.

The majority of the respondents (81.5%) regarded the demonstrations on the central table very useful and the course in terms of general interest as interesting.

Self-pacing, replay of tape and demonstration on the central table were among the most useful features and the wearing of headphones and listening to the tape were most irritating features of the independent study study session as rated by the respondents.

Of the students surveyed, the majority found the study guides very informative and interesting but the self-quizzes and exercises very difficult, lengthy and time-consuming.

The general assembly sessions were not found highly stimulating by approximately two thirds of the student-teachers surveyed.

DISCUSSION

The adoption of the A-T model accomplished the educational objective of exposing the pre-service teachers to a new method of teaching and providing them with valuable educational experiences. The analysis of the data indicated that approximately two-thirds of the student-teachers surveyed "liked" the A-T instructional technique and regarded it a good and effective method. The evaluation information obtained from the survey helped to make certain decisions to continuously alter and improve the future A-T instructional programmes at the Goulburn College of Advanced Education. This continuous process of improvement and the experimental attitude towards the interpretation of assessment data is helping to make sound judgment upon which individually guided science instruction can occur.

Postlethwait's audio-tutorial approach assumes that (1) students learn at different rates and in varied ways; (2) students are capable of being self-propelling; and (3) students can learn independently. It is a multi-media approach in which learning is enhanced by engaging the various senses in receiving stimuli and by maximizing opportunities for the learner to interact with the material and to be in control of his own learning, it is not able to provide individualized instruction in the total sense. To some extent it does provide for the individualization of the rate of learner progress during the independent study session, the learner, however, is required to accomplish the A-T unit within a period of one week. He is not free in the total sense of "individualized rate." Furthermore, in the typical A-T approach the student progresses in a linear fashion through various teacher-structured and pre-determined administrative assignments. He is not taught by an individualized instructional method. As the present study shows, many students do not learn well or do not like to learn using the A-T method. The multi-media A-T approach does not successfully cope with the individual differences in

cognitive style and conceptual complexity encountered in the typical classroom. It provides a learning environment appropriate for some individuals and not for others. It is, therefore, suggested that the A-T approach to instruction should not be the sole method for the transmission of the learning materials of a programme. It should be regarded as an optional instructional technique serving a certain proportion of students. Geisert (1977) suggested the "learning modules" approach, which has all of the components of the A-T method but is broader and more flexible in its approach to individualized instruction.

REFERENCES

BLOOM, B.S., HASTINGS, J.T. & MADUS, G.F. Handbook on formative and summative evaluation of student learning. New York: McGraw Hill, 1971.

EDLING, J.V. Individualized instruction: A manual for administrators. Corvallis: Continuing education publications, Oregon State University, 1971.

GEISERT, P. Individualized Instruction: Can A-T meet the challenge of the future? The American Biology Teacher, 1977, 39, 51.

KHAN, A.G. Effects of audio-tutorial and conventional techniques on cognitive achievements. Journal of Research in Science Teaching, 1979, 16.

NORRIS, R.A., HEIKKINEN, M. & ARMONSTRONG, T. Alternatives for individualised biology: The importance of cognitive style and conceptual complexity. The American Biology Teacher, 1975, 37, 291.

POSTLETHWAIT, S.N., NOVAK, J. & MURRAY, H.T. Jr. The audio-tutorial approach to learning through independent study and integrated experiences. Minneapolis: Burgess, 1969.

Research in Science Education, 1979, 9, 149-157

SCIENCE CLASSROOM PRACTICES AND DESIGN CHARACTERISTICS OF INSTRUCTIONAL MATERIALS: A CASE STUDY

John Dekkers and Anne Dowd

INTRODUCTION

Science teachers and administrators are often faced with the task of making selections from greatly varying instructional materials and then implementing their selection. This is no mean task, for not only has the range of materials increased, but also the degree of sophistication of the instructional design is often such that inservice training is required prior to their use in the classroom. Moreover, whatever selection procedure is used does not assure successful implementation. It has been suggested that successful implementation of a curriculum depends on:

- the extent that the philosophy expressed by the developers has been realised in the instructional materials;

- the principles of instruction;

- teacher understanding of the instructional philosophy;

- teacher training in the use of the materials.
(Ramsey & Howe, 1969; Northfield, 1975; Belanger, 1964).

However, from research that has examined the nature of science classroom practices it has been revealed that:

> an unexciting picture of science classrooms emerges...teacher talk predominates, pupils respond more to the teacher than to other pupils, pupils rarely initiate discussions, little time is given to motivation or to exciting interest in science and probing, extending or exploring ideas and evaluation are short lived (Tisher, 1972, p.39).

It also appears that teachers in general will continue to maintain their own teaching style and make only minor adjustments to teach new curricula (Belanger, 1964; Power & Tisher, 1975; Power, 1977).

If science educators and administrators desire to implement instructional materials in the classroom in keeping with the design requirements and characteristics of the developers, then there appears a need for developers to make known to potential users those attributes that should be fulfilled and principles to be implemented in order to ensure successful implementation of the instructional materials.

This study explores the characteristics of the instructional materials and the nature of classroom practices of two curriculum development projects, namely, Foundational Approaches in Science Teaching (FAST) (1974) and Web of Life course (1973).

FAST is a three year sequential laboratory and field centred "inquiry" science program for grades 8, 9 and 10. Although FAST is not widely used in Australia, it is currently extensively adopted in a number of states in the U.S.A. The broad objectives of the FAST program are to provide students with an understanding of foundational concepts and methodologies in the biological, physical and earth sciences that relate to man's use of the environment (Pottenger, 1974).

The Web of Life course is an "inquiry" oriented biology course for senior high school students. It is an adaptation of the Biological Sciences Curriculum Study (BSCS) and has been extensively used in Australian schools since 1967. The Web of Life course is designed to provide information about current biological knowledge and an understanding of science as a process of inquiry.

METHOD

Instruments Used in the Study.

Science Curriculum Analysis Schedule (SCAS). This instrument provides a systematic and objective method for the analysis of published instructional materials and procedures by a curriculum development project. Its design is based on constructs generally accepted as central to micro and macro curriculum designs and materials analysis systems (Allen, 1974; Eash, 1970; Knight, 1971). SCAS not only provides a data base for critical analysis of a curriculum but also highlights instructional design features.

Activity Checklists. A Biology Activity Checklist and Science Activity Checklist measured the extent of occurrence of particular classroom activities. Items in the activity checklist relate to student and teacher activities in the following categories:

(i) Nature of Science: e.g. We discuss the way developments in science have affected our way of life.

(ii) General Laboratory Procedures: e.g. We discuss conclusions from experiments after each investigation.

(iii) Inquiry Laboratory Procedures: e.g. We have class discussions about unused data from laboratory investigations.

(iv) Evaluation: e.g. We get grades for our laboratory investigations.

(v) Class Activities: e.g. Conclusions from class discussions are made by our teacher.

Students respond to each item on a 5-point temporal scale (almost always, most of the time, often, some of the time, very seldom). The activity checklists thus provide a measure of the level of inquiry orientation in classroom activities. Content validity of SAC has been established in previous studies (Kochendorfer, 1968; Dekkers, 1977).

TABLE 1

SUMMARY OF ANALYSIS OF FAST AND THE WEB OF LIFE COURSE

Construct	FAST	Web of Life
Theory and Rationale		
Curriculum Theory	Based on the curriculum theory developed by King & Brownell (1966)	Not specifically based on a curriculum theory.
Learning Theory	None explicitly stated but there is a strong influence of the "theory" of learning-by-discovery.	None specifically referred to.
Rationale	Factors accounted for in both curricula include the discipline, teacher and society. In addition the FAST curriculum relates extensively to technology and environmental issues.	
Content and Structure		
Content	Based on interdisciplinary utility and social relevancy.	Content intended to be modern and reflect the enterprise of science in general.
	A three tiered spiral program bringing together studies in the sciences, technology and social sciences.	Contains a central core of studies and basis topics which develop essential and major ideas.
Methodology/Materials		
Teacher Competencies	Has a mandatory Teacher Training programme and follow up session essential to fully prepare FAST teachers.	No stated required competencies.
Support Services	Has an on going consultative service and workshops.	Advice provided by A.A.S. on request.
Teaching Strategies	Teacher strategies are very explicitly stated. Role of Teacher is that of a research director and facilitator setting tasks, asking questions and making suggestions for further investigation. Questions and discussion hold a prominent place in all activities and are emphasised at the beginning and end of each investigation.	Actual classroom strategies are not specified although aspects of Teacher strategies are explained. Teacher is described as a leader and director of student activities, however, these are largely teacher directed.
Innovation/Individual Differences	Teacher can be innovative and suggestions are provided on how student materials can be modified for different ages and ability groups. Challenge problems and optional work-sheets are provided.	The curricula is highly structured and there is little opportunity for Teacher to be innovative. Student can be streamed into different ability classes. A basic core of activities can be taught and provision made for the more able student to carry out additional work on their own.
Reading Level of Materials	Reading levels of the materials show them appropriate for 7th grade.	Text is written for average students. Fry and Flesch calculations of readability is consistent with age group.
Evaluation Procedures & Devices		
Learner Evaluation	Both informal and formal evaluation plans and materials are provided. Evaluation materials can be used to identify learning difficulties etc. in cognitive areas and laboratory skills.	No set evaluation plan provided. Diagnostic and end of unit test available. Evaluation programme does not specifically identify learning difficulties.
Teacher Self Evaluation	Possible through consultative visits.	None

Classroom Observation Schedule. Events and behaviours in classes as reported in this study were recorded using a direct observation instrument, the Classroom Observation Schedule (COS). This instrument is developed from observation schedules devised by Evans and Balzer (1970) and Dekkers and Shield (1978). The schedule provides a measure of the extent of student individual work, laboratory work, etc., the development of inquiry processes and an assessment of teacher style.

Sample, Data Collection and Analysis.

The checklists were administered to junior science classes using FAST (N = 312; 9 classes) and Grade 12 biology classes using the Web of Life (N = 302; 19 classes) in Queensland. Classroom observations of a number of science and biology classes were made following the administration of the checklist, each class was observed for a minimum of five successive lessons.

RESULTS

The analysis of both curricula using SCAS reveal a number of similar design characteristics. Both FAST and the Web of Life course source materials contain:

(i) a detailed rationale and overview. In both cases the curriculum is based on the two components of the Disciplinary Theory, that is, the concepts, theories and generalisations and the processes of science. Also, the content of each is modern.

(ii) the assumptions underlying the curriculum and the method(s) of instruction. At a general level these are very similar for both curricula.

(iii) overall strategies of instruction i.e. teacher and student roles.

(iv) details of equipment and material requirements for student activities.

There are, however, a number of distinct differences between both curricula and these are highlighted in Table 1. It can be seen that in contrast to the Web of Life course the FAST curriculum:

- has foundations firmly based on theoretical premises;

- has a very detailed methodology, specifies required teacher competencies and provides support services. Nevertheless, the teacher can still be very innovative;

- provides a more comprehensive evaluation.

Table 2 presents results for student data based on item mean scores in the activity checklist categories. These scores are derived at by dividing the category mean score by the number of items in each activity checklist

TABLE 2

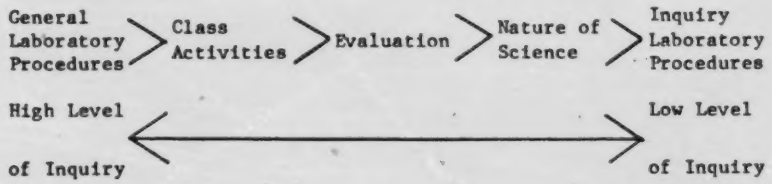
The Level of Inquiry Orientations in FAST and Web of Life Classes

Level of Inquiry Orientation	Item Mean Score	Categories									
		Nature of Science		Laboratory Procedures: General		Laboratory Procedures: Inquiry		Evaluation		Class Activities	
		Web of Life	FAST	Web of Life	FAST	Web of Life	FAST	Web of Life	FAST	Web of Life	FAST
High level of inquiry	2.2-2.4			5,6,7,8, 11,14,17, 18,19	1,2,4,5, 6,7			5			3
	2.2-2.6			4,13,15, 16	3,8,9			6,10		4,6,8,11	1,2,5,6, 9
	2.6-2.8		1,2	3,10,12			1,3,4,5, 9			7,12,14, 16,19	4,7,8
	2.8-3.0		5,6,9,4	1,2			2,6,8	2	3,4,5,8	1,5,13,15 17,18	
	3.0-3.2	5	7,3	9			7	7,8,9,11, 16,17,19	1,2,7, 9	2,9,10	
	3.2-3.4	6,8,11	8			8		3,12,14, 18	6	3	
	3.4-3.6	4,7,19				11		1,4,15			
	3.6-3.8	10,12,14, 15,17,18				5,6,7		13			
	3.8-4.0	2,16				15,17					
	4.0-4.2	1,3,9				14					
Low level of inquiry	4.2-4.4	13				1,2,3,4, 9,10,13, 12,16,18,19					

Numbers in category columns refer to class numbers (FAST, N=9 classes; Web of Life, N=30 classes)

category. A high mean score indicates a high and low level of inquiry practice respectively.

Examination of Table 2 reveals that for all the activity checklist categories FAST classes have lower item mean scores than the Web of Life classes. That is, there is a higher level of inquiry in FAST classes. This is particularly evident for the category Inquiry Laboratory Procedures. Also in contrast to FAST classes, the Web of Life classes vary considerably in the level of inquiry in the different categories. The overall order of categories in terms of level of inquiry orientation for the Web of Life classes is as follows:



For FAST classes this trend is not evident.

Direct observational data using the Classroom Observation Schedule was obtained from only a rather limited number of classes. Despite this, it was evident that there was great variation in the nature of class activities for Web of Life classes even though all classes were working on the same section. This was not the case for FAST classes. In the Web of Life classes for example, teacher dominated content development ranged from 3% to 60% of total class time and student individual work ranged from 20% to 72% of total class time. These results were in agreement with student responses on BAC and the results of earlier studies (Balzer & Evans, 1970; Parakh, 1969). Furthermore in the classes observed, little use was made of the categories closely related to inquiry teaching processes, tentativeness of knowledge, generalisation beyond the actual lesson content, or links with other fields of knowledge.

For FAST in the case of laboratory investigations, teacher dominated talk ranged from 20% to 40% but this tended to be confined to procedural tasks. In other class time, individual student work ranged from 50% to 70% of total class time. Class discussions, particularly those following laboratory investigations, tended to focus on inquiry processes such as the data collection procedures used, interpretation of data, formulation of conclusion and the discussion of unusual results. In these discussions teacher talk ranged from 40% to 60%.

DISCUSSIONS AND CONCLUSIONS

It has been outlined above that the two curricula have a number of similar characteristics. Both are large scale "inquiry" oriented science curriculum development projects. Also, analysis of source materials using SCAS reveals common features in their rationales, strategies of instruction and composition of instructional materials. However, major differences occur between FAST and the Web of Life in the extent that their

instructional and evaluation strategies and teacher support from the developers are made explicit and accommodated for. FAST in contrast to the Web of Life course has a well developed and active teacher training programme and support network. The teacher training programme aims to develop teacher inquiry techniques needed for discussion activities, laboratory and field investigations. It also provides teachers with an understanding of the philosophy and rationale of the programme. The FAST support network includes consultations with the project developers and workshops which provide opportunities for teachers to share experiences, ideas and to discuss teacher and/or learner problems. They also provide an opportunity to gain deeper insights into the programme, its philosophy, objectives and teaching methodology. A further aspect of the teacher training programme is that in teacher training workshops experienced FAST teachers can act as trainers and leaders. This innovative feature acts as a stimulus for further professional teacher development.

Results from both the activity checklist and classroom observation indicate that the level of inquiry orientation in Web of Life classes was relatively low, particularly for the categories Nature of Science and Inquiry Laboratory Procedures. Also in contrast to FAST, there was a greater variance in the nature of classroom practices. It would seem that successful implementation of a science programme requires, or at least enhances, the chance of successful implementation by the following factors:

- the existence of extensively developed instructional strategies,

- the availability of a teacher training programme complimented by a support network.

The need for this latter factor was highlighted in a recent report by Owen and Tisher (1977).

Furthermore, it is also suggested that teacher training and development toward a thorough understanding of the philosophy and objectives of the FAST programme has contributed to its successful implementation. In a related study on the Web of Life course (Dowd, 1979) it was shown that teachers are not well aware of the Web of Life philosophy and principles of instruction. In this study a number of reasons could be put forward to account for this situation. Only two of the teachers in the sample experienced inservice training on the Web of Life course and this could in part account for the varied practices that were being used. Perhaps teacher material is not sufficiently explicit. A curriculum analysis of the Web of Life course did reveal that the developers considered an understanding of the curriculum philosophy and principle of instruction as essential for teaching the course. However, no clear statement of the philosophy and of the structure of the discipline is provided. Nevertheless, the principles of instruction are discussed in the Web of Life Teachers Guide and in other BSCS materials. It does seem that teachers were not using available information to the the full. In part this could be due to the fact that the required teacher information is not very clearly set out.

Whilst the foregoing has focussed on current practices and design characteristics of two programmes, this paper also illustrates how a detailed analysis of the source materials of an instructional package can

be used as an indicator of whether or not a programme can be successfully implemented as described.

REFERENCES

ALLEN, L. Science curriculum inventory. University of Hawaii, 1974.

BELANGER, M. The study and teaching of new curricula. The Science Teacher, 1964, 31(7), 31-36.

DEKKERS, J. Science activity checklist. Unpublished Manuscript, James Cook University of North Queensland, 1977.

DEKKERS, J. & SHIELD, M. An instrument for recording events and behaviours in the mathematics classroom. Research in Mathematics Education in Australia, 1978, 3, 48-54.

DOWD, A. Teacher and student perceptions of the Web of Life course. Unpublished results, James Cook University of North Queensland, 1979.

EASH, J.M. Developing an instrument for assessing instructional materials. Paper presented at American Educational Research Association Annual Conference, Minnesota, 1970.

EVANS, T.P. & BALZER, A.L. An inductive approach to the study of bilogy teacher behaviours. Journal of Research in Science Teaching, 1970, 7, 47-56.

KING, A.R. & BROWNEL, J.A. T curriculum and the disciplines of knowledge: A theory of curriculum practice. New York: John Wiley and Sons, 1966.

KNIGHT, M. Curriculum materials analysis system. Boulder: Social Science Education Consortium Inc., 1971.

KOCHENDORFER, L.H. The development of a student checklist to determine classroom teaching practices in high school biology. In A. Lee (ed.) Research and curriculum development in science education. Texas: Science Education, University of Texas, 1968.

MORGAN, K.P., BEST, E.D., BETJEMAN, K., LARKUM, A., PITMAN, K. & THOMSON, J. Biological science: The web of life, (2nd ed.). Canberra: Australian Academy of Science, 1973.

NORTHFIELD, J.R. Changing perspectives of the science education that may be associated with curriculum development. Research in Science Education, 1975, 5, 23-32.

OWEN, J.M. & TISHER, R.P. Curriculum adoption: The fate of a national curriculum project in Australia. Canberra: Curriculum Development Centre, 1978.

PARAKH, J.S. A study of teacher-pupil interaction in high school biology classes. Journal of Research in Science Teaching, 1969, 6(3), 284-292.

POTTENGER, F. et al. Foundational approaches in science teaching. Hawaii: Curriculum Research and Development Group, University of Hawaii, 1974.

POWER, C.N. A critical review of science classroom interaction studies. Studies in Science Education, 1977, 4, 1-30.

POWER, C.N. & TISHER, R.P. A study of the effects of teaching strategies in ASEP classrooms. Australian Journal of Education, 1975, 19(2), 127-145.

RAMSEY, C.A. & HOWE, R.W. An analysis of research on instructional procedures in secondary school science. Part 1 Outcomes of Instruction. Science Teacher, 1969, 36(3), 62-70.

TISHER, R.P. Classroom interaction analysis: the new religion. Research, 1972, 2, 35-49.

Research in Science Education,1979,9,159-167

CRITICAL THINKING ABILITY IN SCIENCE—WHAT CAN BE MEASURED?

M. J. Sullivan and C. J. Dawson

INTRODUCTION

A previous paper (Sullivan, 1977) described some problems which had arisen in the construction of tests of original thinking ability in science-based situations. The present paper discusses the partial resolution of these problems and presents some findings on the nature and testing of critical thinking ability.

THE MEANING OF "CRITICAL THINKING"

For such an apparently simple term as critical thinking, there is a wide diversity of interpretations. Some writers, e.g. Burton et al. (1960), Crow & Crow (1954) have equated critical thinking with Dewey's reflective thinking, which consists of:

> active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and further conclusions to which it tends (Dewey, 1933, p.9).

Later, Dewey introduced the idea that reflective thinking involved "doubt and perplexity" and an "act of searching" (Dewey, 1933, p.12) which connected reflective thinking with problem solving, and many writers subsequently have equated problem-solving with critical thinking (Organ,1965; Poel,1970).

The term "critical thinking" began to appear frequently in the literature of education and psychology in the late thirties; no doubt a connection existed with the interests of educators at this time with propaganda. Glaser in 1941 described critical thinking in terms of three factors:

> an attitude of inquiry; knowledge and understanding of the methods of inquiry and logic; skills in the application of such methods (Glaser,1941).

Many psychologists and educators have agreed more or less with this concept, but there is a widely held opinion that incorporated into the knowledge component should be a knowledge of the structure of the subject area involved (Ennis, 1963; Russell, 1956; Benson & Young, 1971; Ausubel, 1968). Ennis asserts that "some are good critical thinkers in one area and not in other areas" while Ausubel dismisses the idea of development of general critical thinking ability by stating that "it can never amount to more than a critical approach to the teaching of particular subject-matter disciplines."

As an outcome of instruction, critical thinking ability has received much interest and support: "Tests in Print II" (Buros, 1974) cites 144 references to use of the Watson-Glaser Critical Thinking Appraisal (WGCTA) up to 1972, and many of these studies were in science education. In most cases, however, the researcher's definition of critical thinking was not given and the instrument was used without much attention to its validity in the particular context. The WGCTA covers a wide range of historical, social and hypothetical situations, with very few items connected with science: thus for studies which involve measurements of change of critical thinking ability as a result of studying science courses, the WGCTA has, according to Ennis and others, limited usefulness.

So, as we wished to study the critical thinking ability of students of a matriculation chemistry course, we needed to provide a suitable instrument. Having noted the problem of definition, it was necessary to have a clear and preferably a concise one to work from. Ennis's (1962) simple definition of critical thinking as "the correct assessment of statements" has considerable appeal, and can be considered as implying all three of Glaser's factors.

The WGCTA does in fact require the correct assessment of statements, which together with its lasting popularity after many years of development, led us to conclude that we were unlikely to improve upon its format. We constructed all new items based largely on chemistry situations (Form C); subsequently a largely physics form (Form P) was developed. The problems which we had with low internal consistencies and poor inter-form correlations were described in the previous paper (Sullivan, 1977).

A TEST OF CRITICAL THINKING IN PHYSICAL SCIENCES

A suitable starting-point for the description of our more recent findings would be a re-statement of the more puzzling results obtained previously. Although the initially low internal consistencies for the two forms had been improved by the usual methods, some poor inter-test and inter-subtest correlations remained. Thus both forms had internal consistency coefficients (α) of 0.67, but the correlation between them (Pearson's r) was only 0.36. Correlations between corresponding subtests in the two forms varied from around zero (Evaluation of Arguments, Deductions) to about 0.5 (Recognition of Assumptions).

Conventional methods of test improvement have since been employed, maintaining the subtest structure: the result of this work has been to gradually remove most of the anomolies previously encountered.

It has been assumed, both by Watson and Glaser and by ourselves, that only one ability is being measured: the different types of item (subtests) are used only to introduce variety into the criteria on which statements are to be assessed. Ennis (1963) lists 12 ways in which a statement might be assessed: the WGCTA does not cover all of them. Two new subtests were developed, one requiring detection of ambiguity in a line of reasoning and the other a decision as to whether a statement was specific enough. The ambiguity subtest was not successful: the items seemed to be so difficult that most respondents were guessing. More work needs to be done on

specificity subtests, but they appear to be a useful addition to the range.

Eventually, five forms of a Test of Critical Thinking in Physical Science (TCTPS) were assembled, having the following characteristics:

Form	Items	Mean	S.D.	Reliability (α)	Correlations(r) Form C	Form A
C	50	29.05	6.10	.76		
P	50	30.81	6.83	.76	.66	
A	50	29.77	6.78	.78		
B	50	31.32	6.12	.75		.77
S	100	60.97	11.33	.89		

(N = 220 approx.)

Form C has items largely, but not entirely, with a chemistry base, and form P has largely physics items. A and B are alternate forms concerned with both chemistry and physics. Form S is a composite whose internal consistency is gratifyingly high. The correlation between Forms C and P is now reasonably high, but is lower than internal consistencies; this can be taken to imply that the subject area does play some part. Items perform slightly differently in the different forms, and those giving the best total internal consistency have been included in each form: a total of 110 items is involved.

Most of the items are essentially true/false, which of course creates a tendency toward low reliability, and it seems that the reliability for the 50-item tests will not easily be pushed above 0.8. However, there is evidence from our observations that students tend to lose some interest and, perhaps, concentration if many questions of the same type appear, that is, if the subtests are too long. The 50-item tests may therefore be more satisfactory if the incentive to complete them is not high.

Some Subtest Characteristics.

It was pointed out earlier that some subtests tended to intercorrelate less well than others, and to have more apparently weak items, so more work was required to bring their properties into line. Interpretation items for example, seem to be easy to construct, but Evaluation of Arguments presents problems. This seems to be true even for the existing subtests in the WGCTA, Forms Ym and Zm, (Watson & Glaser, 1964, pp. 14,15). Deduction and Evaluation subtests have lower internal consistencies on the whole, and correlations tend to be lower. Data obtained by Westbrook & Sellers (1967) and Troxel & Snider (1970) are shown in Table 1. Again lower figures are often associated with the Evaluation subtest.

Some figures for the TCTPS are shown in Tables 2 and 3. Apart from the specificity subtests, which have had less work on them than the others, the evaluation subtests, which from the start had a considerably higher item turnover than average, again show lower internal consistencies than others. The implications of this will be discussed later.

TABLE 1(a)

Correlations of WGCTA Scores with QW* and HN** Scores

Test	1	2	3	4	5	6	7
1. WG (Inf.)							
2. WG (Ass.)	.27						
3. WG (Ded.)	.34	.34					
4. WG (Int.)	.35	.27	.52				
5. WG (Eval.)	.26	.12	.29	.30			
6. WG (Total)	.65	.64	.77	.73	.52		
7. QW	.30	.25	.37	.37	.17	.44	
8. HN	.35	.25	.43	.43	.30	.53	.57

* Quick-word Test
** Henmon-Nelson Tests of Mental Ability

TABLE 1(b)

WGCTA, TOUS and ACS Chemistry Intercorrelations

Test	2	3	4	5	6	7	8
1. ACS Chemistry*	0.46	0.23	0.13	0.37	0.29	0.22	0.38
2. TOUS**		0.37	0.23	0.42	0.35	0.32	0.52
3. WG (Inf.)			0.17	0.35	0.22	0.20	0.56
4. WG (Ass.)				0.26	0.20	0.12	0.56
5. WG (Ded.)					0.47	0.35	0.77
6. WG (Int.)						0.40	0.75
7. WG (Eval.)							0.59
8. WG (Total)							

* ACS Co-operative Chemistry Examination
** Test on Understanding Science

TABLE 2

Subtest Means and Reliabilities, Form S.

Subtest	No. of Items	Mean	S.D.	Reliability (α)
1. Inference	19	8.88	2.96	0.52
2. Recognition of Assumptions	15	10.44	2.59	0.62
3. Deduction	17	10.43	2.97	0.52
4. Interpretation	18	11.34	3.14	0.60
5. Evaluation of Arguments	18	13.30	3.51	0.41
6. Specificity of Statements	13	6.58	2.79	0.52

WHAT DO CRITICAL THINKING TEST MEASURE?

A number of workers have applied factor analytic techniques to critical thinking tests, including the WGCTA, the American Council on Education's "Test of Critical Thinking" (American Council, 1951) and Enniss & Millman's "Cornell Critical Thinking Test". Not surprisingly, since any intellectual process is complex, many factors have been identified. Rust (1960,1962) identified three from the many as:

1. General Reasoning
2. Logical Discrimination
3. Verbal Reasoning (Semantics).

Westbrook & Sellers (1967) also detected a general factor, but Follman (1970) did not, finding a larger number of factors, the one accounting for the largest percentage of the variance being "judging if conclusions follow from evidence". Preliminary analysis of the TCTPS revealed 17 factors with eigen-values greater than 1, but none accounted for more than 6 per cent of the variance.

As factor analysis was not very helpful, we have had to be content with a less rigorous description of our tests, relying on face validity, internal consistencies, and correlations. Scores on a shorter version of form C correlated rather poorly with a test of chemical knowledge and with the South Australian matriculation examination (r about 0.3), and a little better with tests on understanding science and problem solving skills (r about 0.5). Correlations with the WGCTA (from earlier versions, about 0.6) indicate that there is much common ground, but that our tests do contain factors connected with their physical science base. However, this is not

TABLE 3

Subtest Intercorrelations (Pearson's r)

Subtest	1	2	3	4	5	6	13	14
1. C-Inference							.72	
2. C-Assumptions	.36						.63	
3. C-Deduction	.39	.33					.73	
4. C-Interpretation	.56	.55	.50				.70	
5. C-Evaluation	.36	.20	.30	.55			.53	
6. C-Specificity	.15	.23	.19	-	.20		.53	
7. P-Inference	.39							.67
8. P-Assumptions		.44						.70
9. P-Deduction			.30					.74
10. P-Interpretation				.40				.70
11. P-Evaluation					.43			.47
12. P-Specificity						.40		.53
13. C-Total								
14. P-Total								

	8	9	10	11	12	16	17	18	19	20	21
7. P-Inference	.47	.41	.33	.27	.11						
8. P-Assumptions		.41	.30	.33	.19						
9. P-Deduction			.29	.28	.25						
10. P-Interpretation				.26	-						
11. P-Evaluation					.11						
12. P-Specificity											
15. S-Inference						.48	.53	.61	.42	.18	.81
16. S-Assumptions							.40	.47	.30	.22	.74
17. S-Deduction								.52	.46	.34	.80
18. S-Interpretation									.55	-	.83
19. S-Evaluation										.23	.71
20. S-Specificity											.59
21. S-Total											

the kind of factual knowledge measured by the usual subject tests.

It seems, therefore, that the knowledge components of critical thinking, both of logic and subject structure, and the skills involved in applying such knowledge, can be incorporated into reliable and apparently valid tests. The problem of measuring a questioning attitude, however, is a much more difficult one. As soon as a person is invited to assess statements, any attempt to assess his habit of doing so is negated. Quite different techniques are required to assess critical habits, and perhaps Milgram's work (1974) on compliance with authority can give some insight into the area. This is an important aspect of critical thinking, but its treatment may require considerasble skills in experimental psychology.

The problems already outlined in connection with evaluation of arguments require some explanation. A plausible one is not hard to find. The items showing poor discrimination were mostly connected with controversial subjects, e.g. nuclear power, chemical warfare: in spite of instructions to consider only the evidence supplied, able students with knowledge of, and presumably, views on these subjects could not resist answering in accordance with their views. Thus the argument (against closing down all chemical plants which emit waste) : "Some pollutants cause no permanent damage but are difficult to remove" is a strong one if accepted as true (which respondents were instructed to do, whether or not they personally believed it to be so). However, its discrimination index was negative, suggesting that the normally good critical thinkers had a horror of the emotive word pollutants which caused them to reject any argument which suggests that these could ever be tolerated (which, of course, they inevitably are).

A number of other items were rejected for poor discrimination, usually accompanying low facilities; reluctantly, because they undoubtedly tested a person's ability to view evidence unemotionally, surely an essential part of critical thinking. The evalutaion items remaining were thus depleted of controversial topics, which seems unfortunate. There is a case for separating such items from the main body, and creating a new scale incorporating an "emotion" factor, but it is impossible to predict the actual involvement of feelings in a science situation, and their degree and direction will not be the same for all. This is a problem requiring more investigation.

Thus critical thinking is a complex subject and its testing in areas involving more than pure knowledge or logic is difficult. Perhaps it is unrealistic to expect people to adopt highly consistent attitudes; society would be very different if they did.

REFERENCES

AMERICAN COUNCIL ON EDUCATION. A test of critical thinking, Form G :Instructor's manual. Washington: The Council, 1951.

AUSUBEL, D.P. Educational psychology: A cognitive view. New York: Holt, Rinehart and Winston, 1968.

BENSON, B.W., & YOUNG, L.L. The development and implementation of an instrument to assess cognitive performance in high school biology. Journal of Research in Science Teaching, 1971, 8, 211-224.

BUROS, O.K. (ed.). Tests in print, II. New Jersey: Gryphon, 1974.

BURTON, W.H., KIMBALL, R.B. & WING, R.L. Education for effective thinking. New York: Appleton-Century Crofts, 1960.

CROW, L.D. & CROW, A. Educational psychology. New York: American Book, 1954.

DEWEY, J. How we think (rev. ed.). Boston: Heath, 1933.

ENNIS, R.H. A concept of critical thinking. Harvard Educational Review, 1962, 32, 81-111.

ENNIS, R.H. Needed--research in critical thinking. Educational Leadership, 1963, 21, 17-39.

ENNIS, R.H. & MILLMAN, J. The Cornell critical thinking test. University of Illinois Critical Thinking Project, 1971.

FOLLMAN, J. et al. Factor analysis of critical thinking, logical reasoning and English subtests. Journal of Experimental Education, 1970, 38(4), 11-16.

GLASER, E.M. An experiment in the development of critical thinking. Contributions to Education No. 843. New York: Columbia University, 1941.

MILGRAM, S. Obedience to Authority. London: Tavistock, 1974.

MORSE, H.T. & McCUNE, G.H. Selected items for the testing of study skills. Bulletin No. 15. National Council for the Social Studies.

ORGAN, T.W. The art of critical thinking. Boston: Houghton Mifflin, 1965.

POEL, R.H. Critical thinking as related to PSSC and non-PSSC physics programs. Ph.D. Thesis, Western Michigan University, 1970.

RUSSELL, D.H. Children's thinking. Boston: Ginn, 1956.

RUST, V. Factor analysis of three tests of critical thinking. Journal of Experimental Education, 1960, 29, 177-182.

RUST, V., JONES, R.S. & KAISER, H.F. A factor analytic study of critical thinking. Journal of Educational Research, 1962, 55. 253-9.

SULLIVAN, M.J. Critical thinking in science-some problems in measurement. Research in Science Education, 1977, 7, 131-8.

TROXEL, V.A., & SNIDER, C.F. Correlations among student outcomes on the Test on Understanding Science, Watson-Glaser Critical Thinking Appraisal and the ACS Cooperative Examination--General Chemistry. School

Science and Mathematics, 1970, 70, 73-6.

WATSON, G. & GLASER, E.M. Watson-Glaser critical thinking appraisal, Form Ym: Manual, New York: Harcourt, Brace and World, 1964.

WESTBROOK, B.W. & SELLERS, J.R. Critical thinking, intelligence and vocabulary. Educational and Psychological Measurement, 1967, 27, 443-6.

Research in Science Education,1979,9,169-175

COGNITIVE CONFLICT: ITS NATURE AND USE IN THE TEACHING OF SCIENCE

J. A. Rowell and C. J. Dawson

INTRODUCTION

Many theories which are currently, or potentially, important to educational practice are committed to cognitive conflict as a basic component of the learning process e.g. Bruner (1966); Klahr and Wallace (1976); Case (1978); Piaget, (see Moessinger, 1977); Inhelder, Sinclair and Bovet (1974); Karmiloff-Smith and Inhelder (1974). However, use of the same terminology tends to obscure the fact that theories differ in the interpretation which they impose on that concept. The purpose of this paper is to comment briefly on some of those interpretations and to examine the educational practices which follow from theoretical commitment.

Conflict and Its Resolution in the Work of Bruner.

A fundamental assumption underlying Bruner's developmental theory is that knowledge can be represented internally in three unique, but partially translatable systems; enactive, ikonic and symbolic (Bruner, 1966). Further, it is posited that the order of emergence of each of these forms of representation is fixed and that each is typical of a distinct developmental stage, though, once emerged, each system continues its unique existence.

Within this theory cognitive conflict occurs when there is a mismatch between information encoded in two of the representational systems, between "...what one sees and how one says it, or how one must act overtly and how the world appears." (Bruner et al. 1966, p. 11). Resolution of the contradition results in mental growth.

The screening experiment of Frank (Bruner et al. 1966, Ch. 9) is quoted as empirical support for Bruner's notion of conflict; though Strauss (1972) contests the value of the Frank results. However, the educational translations of the idea seem very sparse. Even in Bruner's relatively early work, published at the same time as the development of his analysis of conflict, it is mutual support rather than conflict which is in evidence (Bruner & Kenny, 1965).

Since that time a change in emphasis has occurred within the theory toward a reinterpretation of Bruner's modes of representation as modes of instruction i.e. as modes differing in the skills which they assume and/or develop in the learner but with each providing the same knowledge (Olson, 1972). Bruner himself (Olson & Bruner, 1974), perhaps predictably, retains his original notion of representational modes which he relates to the three forms of instruction.

Reinterpretation as instructional forms removes some fundamental criticisms of the theory (McNeil, 1967; Fodor, 1976). In its emphasis on the acquisition of the skills involved in extracting and utilizing knowledge in these different forms it draws attention also to a wider, more inclusive view of conflict. Now conflict and difficulty in problem solving, together with failed communication, are seen as elicitors of "deuteropraxis" or second order information processing; a reorganization and integration of specific experiences which, if successful, makes the problem comprehensible to the individual and potentially teachable to others (Olson & Bruner 1974, pp. 142-148). Bruner's views, however, are themselves a deuteropraxic account which, to paraphrase his comments on such matters (Olson & Bruner 1974, p. 147), may be a powerful summary of experience but also may be woefully inadequate as an alternative to it. As a more specific guide to that experience we turn therefore to other authors for practical illumination.

Conflict and Its Resolution in the Work of Case.

In a series of articles Case (1975, 1978) has developed an information-processing view of cognitive development with some similarities to that of Bruner and has provided detailed, specific educational translations of the general principles involved, including the concept of cognitive conflict.

Briefly, Case has attempted to build a bridge between the behaviourally based ideas of Gagné (1970) and the developmentalist theory of Piaget (Piaget & Inhelder, 1969). His instructional technology (Case, 1978) incorporates what we have labelled an interactionist view of conflict (Rowell & Dawson, 1979) i.e. a view of conflict in which the elements (mental strategies and environmental events in this instance) are considered separable, independent objective entities, each with its own structure, which interact. In practice conflict within the Case technology works as follows: Following the use of techniques designed to reveal the learner's spontaneous strategy in relation to the problem posed the next step is to reveal its limitations to him. This is effected by first providing the student with a meaningful procedure for making judgements and then exposing the inadequacies in his spontaneous strategy by carefully contrived counter examples. Once the student recognizes the limitations in his approach - and within the theory there is no reason why he should not - he may then develop his own new strategy, or if this is not possible, he is ready to accept one whose adequacy can be demonstrated to him.

Conflict and Its Resolution in the Work of Piaget.

For the Piagetian, analysis of the knower-known relationship must start from the fundamental organismic-environmental matrix (Furth, 1969, p. 176); a viewpoint described by Lerner (1978) as a "...dual source interdependent interaction effect." What this means is that environmental events are inseparable from the operative structures of the learner; in other words his theory, method of judgement of reality and the reality which he constructs are interdependent and, contrary to the interpretation of Case, he will not necessarily interpret an event as contradictory even though the teacher "knows" what it "obviously" i.e., logically, is.

Piagetian protocols provide a plentiful source of evidence for this interpretation; and no doubt many teachers can quote frustrated attempts at instruction too. As an example consider the ad hoc modifications and inconsistencies of student BAR in relation to the law of floating bodies (Inhelder & Piaget, 1958, p. 33). These occur even though, of course, BAR has available to him a familiar indicator of success - whether the object does float or sink.

The teaching problem is to make actual to the individual his virtual i.e. unappreciated, contradictions. The Case technology assumes that knowledge of a mechanism for judgement provides the necessary and sufficient linkage. However, for Piagetians this makes the learning situation less problematic than they believe it is, as discussed in part above. Even in the situation where counter examples are recognized as such i.e. virtual contradictions become actual, some problems may remain. According to Inhelder (Karmiloff-Smith & Inhelder, 1974), recognition of a contradiction is most likely to occur when the prediction made, which may not be verbal, "...emanates from an already powerful theory expressed in action." What tends to happen next however is not abandonment of that theory coupled with a search for a more adequate one. Rather, if interest remains (Bower, 1974, p. 239), the old theory is retained for as long as possible - being used in situations where it holds true - and a new, independent theory is constructed for the most obvious exceptions, as the learner interprets these. Only later are events unified under a single broader theory.

The question raised is the extent to which instruction should follow or can short circuit this developmental progression. For some a thoughtful methodology for short-circuiting, such as that of Case (1978), does appear to be successful, judged by some criteria. For others problems still occur, which may be related to the "natural" progression observed by Inhelder.

The Piagetian framework emphasises more fully than that of Case a developmental progression in the production of virtual contradictions and of subsequently making these contradictions actual. Within this viewpoint instruction is concerned, where possible, with the mental structures of the individual by making him aware of his virtual contradictions through his attempting to assimilate, i.e., give meaning to, a situation through the dual actions of two uncoordinated structures which lead to different inferences. (Inhelder, Sinclair & Bovet, 1974, p. 249; Bower, 1974, pp. 238-239). It seems to us that three forms of conflict resolution are recognizable. In the first the conflict may be resolved by the logical force, to the individual, of one of the conflicting structures (type 1 resolution). In the second, an appeal is made to a demonstration which the individual interprets as supporting an inference required by his application of one structure but not the other (type 2 resolution). A third form of resolution is evident when neither of two structures is seen as providing a total answer to some experience interpretable through both. In this instance resolution is achieved through the individual's construction of a single unifying idea (type 3 resolution). We offer below illustrations of the first two of these forms of conflict resolution. An illustration of the third is provided by Karmiloff-Smith and Inhelder (1974).

Consider first the problem of volume conservation. Empirical investigations indicating a lack of conservation of this aspect of matter by individuals of all ages are too numerous to detail so perhaps we may be excused for referring to our own studies. We found no less than 22 (67 percent) of 33 year 9 students had problems of varying nature (Rowell & Dawson, 1977) and that even post-graduate university students could be in error (Rowell & Renner, 1976). In an attempt to overcome these problems among students in years 8 and 9 we proceeded as follows:

(a) From the responses of students to group tests of conservation (Rowell & Renner, 1976), two possible pathways to an understanding of volume (as interior and displacement volume) were hypothesized. These included approximately eighty percent of the responses of conservers of weight who had problems with volume. One group of students (Group A) conserved neither interior nor displacement volume but appeared to equate them i.e. to have a unified volume concept. For this group, then, the pathway was direct. If taught the conservation of one "aspect" of volume they should autonatically conserve the other. A second group of students (Group B) did not conserve interior volume but clearly separated this from displacement volume which was conserved on a weight argument. For this group the path appeared less direct. If taught the conservation of interior volume it could be hypothesized that they would not automatically change their weight argument for the conservation of displacement volume, but would do so, and would unify their volume concept, only if made aware of the virtual contradiction inherent in the bases of their conservation explanations.

(b) Using two small groups of four students A and B, representative of groups A and B respectively, separate attempts were made to teach each group the conservation of interior volume by making individuals aware of the contradiction inherent in judging that concept by employing both a number conservation scheme (i.e. same number of same volume objects = same volume) and a scheme of ordinal reasoning applied to a single dimension of a display (e.g. longer = more volume), (type 1 resolution - see above). The relatively simple situation posed by discontinuous matter was tackled first, using one centimetre cubes of metal, and extended to continuous matter (plasticine) once students had "...s heuristic conflict model at their disposal." (Inhelder, Sinclair & Bovet, 1974, p. 249).

(c) As hypothesized, students in group A now automatically conserved displacement volume using the argument of the independence of space occupied and shape. Also as hypothesized students in B did not initially do this. When confronted by the "usual" displacement volume conservation questions, involving identical masses of plasticine of different shape immersed in identical beakers filled to an equal depth with water (Rowell & Renner, 1976), the contradiction in their explanatory schemes remained virtual i.e. unappreciated. Students were unaware of the contradiction because in this instance both their current and potential explanatory schemes lead to the same inference at the level of the question asked (i.e. whether the rise in water level would be the same in each beaker or greater in a particular beaker). These schemes are as follows:

Scheme 1. Interior volume ≠ Displacement volume. Water displaced is determined by the weight of the displacing object. Volume of water

displaced is same when the shape changes because weight and shape are independent.

Scheme 2. Interior volume = Displacement volume. Volume of water displaced is the same when the shape changes because volume and shape are independent.

However, when faced with predicting the changes in water level which would accompany two cubes of equal volume but markedly different weight conflict became actual for some students, judging from their manifest uncertainty and changes of mind. Presumably a physical demonstration at this point could have resolved the issue (type 2 resolution - see above). However, the authors chose instead to present a verbal cue as follows: "Think about when they (the cubes) go in (to the measuring cylinder) they move water out of the way. What do you think might..." Its effect was the reverse to that intended. Before the sentence was finished all four students in B interpreted it as a confirmation of their weight theory - an illustration of the power of "older" beliefs. And all now predicted an equal rise in water levels would be produced by the separate immersion of two cubes of equal weight but differing in volume by a factor of 3.5. Actual immersion, however, resolved the issue, B students immediately recognized the confirmation of their available interior = displacement volume scheme with conservation based on an independence of volume and shape (type 2 resoltuion).

As a second practical illustration of a Piagetian interpretation of cognitive conflict consider the following proposed methodology of counteracting rather common student errors apparent in chemical calculations from equations.

In a recent investigation of teaching the mole concept to Year 11 students we found a number of them made errors either through use of an incorrect stoichiometric ratio or through "revisiting" the equation, i.e. using the equation, or some aspect of it, more than once in the calculation. For example, in answer to "What mass of potassium chlorate would be needed to produce 5.6 dm^3 of oxygen gas at stp?" one student, typifying this error category, worked as follows:

GIVEN $2KClO_3 = 2KCl + 3O_2$

STEP 1 $\therefore$ $2/3 \longrightarrow 1$ mole

STEP 2 $\therefore$ 163.3 g $\longrightarrow 22.4\ dm^3$

STEP 3 $\therefore$ $\frac{163.3 \times 5.6}{22.4} \longrightarrow 5.6\ dm^3$

$= 40.8$ g.

The equation was "revisited" in step 2 in the determination of the mass of potassium chlorate, though not the volume of oxygen; 163.3g is two-thirds of two moles of potassium chlorate.

To overcome these problems we suggested anticipating them by a revision of our methodology which would bring into conflict specific erroneous thought processes and inferences with those of mass and number conservation. For example, consider a reaction represented by

$$2AB_3 = 2A + 3B_2$$

in which A and B symbolise atoms of two elements which can be represented in a concrete model by different sized (and differently coloured to maintain identity) plasticine balls having masses of, say 6 and 1 grams respectively. Suppose the question is asked, "What masses of A and B_2 would be produced by complete decomposition of 36 grams of AB_3?" As the teacher knows the common errors he is able at each step in the calculation to draw attention to the possible conflicts between them, and hence the mental structures presumed responsible, and the inferences derived from employing number and mass conservation schemes, which are made explicit. At each step the conflict may be resolved, as it was for the explanation of displacement volume conservation, by a practical demonstration i.e. following the steps quantitatively on a balance (type 2 resolution). For those students for whom awareness of the conflict has already triggered a type 1 resolution a practical demonstration will confirm their new beliefs.

SUMMARY

The thrust of this article is that although different theories may share some of the same terminology they may differ considerably in the interpretation which they impose on it and, therefore, on the educational practices which they underwrite. Specifically a brief examination has been made here of the concept of cognitive conflict as it occurs in the theories of Bruner, Case and Piaget-Inhelder, and of the similarities and differences in classroom practices which arise from theoretical commitment.

REFERENCES

BOWER, T.G.R. Development in infancy. San Francisco: W.H. Freeman, 1974.

BRUNER, J.S. & KENNY, H.J. Representation and mathematics learning. Monographs of the Society for Research in Child Development, 1965, 30, 50-59.

BRUNER, J.S., OLVER, R.R. & GREENFIELD, P.M. Studies in Cognitive Growth. New York: John Wiley & Sons, 1966.

CASE, R. Gearing the demands of instruction to the developmental capacities of the learner. Review of Educational Research, 1975, 45, 59-87.

CASE, R. A developmentally based theory and technology of instruction. Review of Educational Research, 1978, 48, 439-463.

FODOR, J.A. The language of thought. Sussex: Harvester Press, 1976.

FURTH, H.G. Piaget and knowledge: Theoretical foundations. New Jersey: Prentice-Hall, 1969.

GAGNÉ, R.M. The conditions of learning. New York: Holt, Rinehart & Winston, 1970.

INHELDER, B. & PIAGET, J. The growth of logical thinking from childhood to adolescence. London: Routledge & Kegan Paul, 1958.

INHELDER, B., SINCLAIR, H. & BOVET, M. Learning and the development of cognition. Cambridge, Mass.: Harvard University Press, 1974.

KARMILOFF-SMITH, A. & INHELDER, B. If you want to get ahead get a theory. Cognition, 1974, 3, 195-212.

KLAHR, D. & WALLACE, J.G. Cognitive development: An information-processing view. New Jersey: Lawrence Erlbaum, 1976.

LERNER, R.M. Nature, nurture and dynamic interactionism. Human Development, 1978, 21, 1-20.

McNEIL, D. Cognitions and conservations: two views. Contemporary Psychology, 1967, 12, 530-532.

MOESSINGER, P. Piaget on contradiction. Human Development, 1977, 20, 178-184.

OLSON, D.R. On a theory of instruction: why different forms of instruction result in similar knowledge. Interchange, 1972, 3, 9-24.

OLSON, D.R. & BRUNER, J.S. Learning through experience and learning through media. In Olson, D. (ed.) National Society for the Study of Education 73rd Yearbook, Part 1. Chicago: University of Chicago Press, 1974.

PIAGET, J. & INHELDER, B. The psychology of the child. London: Routledge & Kegan Paul, 1969.

ROWELL, J.A. & DAWSON, C.J. Teaching about floating and sinking: further studies toward closing the gap between cognitive psychology and classroom practice. Science Education, 1977, 61, 527-540.

ROWELL, J.A. & DAWSON, C.J. Skill integration, Piaget and education. Studies in Science Education, 1979, 6, 45-68.

ROWELL, J.A. & RENNER, V.J. Quantity conceptions in university students: another look. British Journal of Psychology, 1976, 67, 1-10.

STRAUSS, S. Inducing cognitive development and learning: a review of short-term training experiments. 1. The organismic developmental approach. Cognition, 1972, 1, 329-357.

Research in Science Education,1979,9, 177-182

CREATIVE AND LOGICAL THINKING IN ADOLESCENTS

Denis Goodrum

INTRODUCTION

If you ask a teacher to list the purposes of schooling, somewhere in the list will more than likely be the statement "To develop children's thinking." Such an aspiration is no doubt worthy, but it is surprising how little we know about the thinking process and how even less, we know about the way thinking can be developed.

Two Modes of Thinking.

Recent evidence suggests thinking can be classified in two distinct modes. One mode, which is believed to be related to the "left hemisphere" of the brain, is predominately involved with analytic logical thinking, especially in verbal and mathematical functions. Its mode of operation is primarily linear and sequential. The thinking operations as conceived by Piaget seem to be related to this type of thinking. Such thinking is often perceived as scientific thinking. It is that which is akin to the rational approach to problem solving. This approach is seen in the majority of our school science courses.

Maslow (1962), Bruner (1966), Adams (1974), Koestler (1967), and Orstein (1972) believe there is another mode of thinking which is also important to the scientist. The other mode, commonly referred as "right hemisperic thinking," is more concerned with holistic activities. Experiments from brain damaged studies, split brain studies and brain studies involving normal subjects, have substantiated the existence of such a thinking mode (Dimond & Beaumont, 1974). These experiments have also delineated the types of thinking associated with this right hemispheric mode. Such activities are primarily concerned with the visual area like spacial relations, spacial orientation, and the recognition of faces. Some studies also indicate imagination, artistic endeavour, musical aptitude and creative-associate thinking is right hemispheric in nature. This mode of thinking processes information more diffusely than does that related to the left hemisphere and demands a ready integration of many inputs at once. While the left hemispheric mode is mainly linear and sequential, the right takes a large amount of information and by considering the information as a whole comes to a conclusion.

Creative Thinking.

Creative thinking is commonly associated with right hemispheric thinking though the relationship is not fully understood. In fact, the concept of creativity has proved to be a difficult one to clearly delineate and define. Freeman et al. state "there is as yet no unified

psychological theory of creativity available to the research worker or the educational practitioner" (Freeman, Butcher & Christie, 1971, p. 2). Despite this lack of clarity regards creativity, the research into the area can be classified into the following three ways:

(a) assessing the creative responses to psychological testing and problem solving. Studies of this type often involve the general theory of human intellect;

(b) identifying and measuring the personality characteristics of individuals and groups considered to be outstandingly creative;

(c) investigating educational methods conducive to the development of creative talent.

The study to be outlined in this paper develops the method used in the first of these areas. Within this method the relationship between creativity and say, general intelligence, is examined by constructing a battery of tests designed to assess creativity and another set to measure intelligence. The correlations within the battery and within the tests are then examined.

Past research has more commonly revealed high correlations between tests of intelligence, low correlations between tests of creativity and correlations between the two kinds of tests that have varied from zero to high. Studies that have suggested a distinct difference between creativity and intelligence include Getzels and Jackson (1962), Wallach and Kogan (1965), Torrance (1960) and Cropley and Maslany (1969). Intelligence in this sense has been mainly viewed in terms of the verbal and mathematical rational mode of thinking as characterized by the left hemisphere.

An interesting theory, first proposed by Torrance, relates to the level of intelligence and creativity. The threshold concept, as it is called, proposes that a minimum level of ability is needed before one can distinguish between creativity and intelligence. Beyond this minimum ability there is little relationship between general ability and the ability to think creatively. Studies by Barker Lunn (1970) and Byrne and Freeman (1971) support this idea.

THE STUDY

The purpose of this paper is to examine a study which investigates the development of creative and rational thinking modes in a number of grade 8, 9 and 10 advanced science classes. This study was part of a much wider study which was involved with determining the nature and relationship of different modes of thinking in adolescents (Goodrum, 1978).

The problem was to see how intelligent science students developed the following types of thinking through the junior high years:

1. Creative thinking,

2. Right hemispheric thinking,

3. Left hemispheric thinking,

4. Piagetian development.

In the study grade 8, 9 and 10 advanced science students from three senior high schools in Perth were tested with a battery of pencil and paper tests. The tests used are as indicated in Table 1.

TABLE I

Tests used in Study

	Thinking type	Test on which Thinking type is based
1.	Verbal Creativity	Alternate Uses
2.	Non-verbal Creativity	Making Objects
3.	Creativity Sum	Sum of 1 and 2
4.	Spacial Relations	Spacial Relations P.M.A.
5.	Spacial Orientation	Spacial Orientation
6.	Right Hemispheric Sum	Sum of 5 and 6
7.	Number Facility	Number Facility P.M.A.
8.	Reasoning	Reasoning P.M.A.
9.	Left Hemispheric Sum	Sum of 7 and 8
10.	Piaget Development	Burney's Test

The left hemisphere specializes in the thinking processes of reasoning and number facility and these two factors are used as a measure of left hemispheric thinking. Spatial relations and spatial orientation are considered to be dominant in the right hemisphere and were selected as representing right hemispheric thinking. Two measures of creativity were used - a non-verbal measure and a verbal measure.

Piagetian cognitive development was measured by Burney's test (1974) which is an extension of a test devised by Tisher.

The total number of students used in the study was 257 in which 83 students were from school "A", 92 students were from school "B" and 82 students were from school "C". In the statistical analysis which employed the analysis of variance statistic, the three schools were treated as separate populations. All raw scores were standardised within each

particular school.

RESULTS

Creativity vs. Grade.

From Table 2 one can see there was no significant relationship between the academic grade of the student and creativity. Such data questions the development of creative thinking of the advanced science students in these three schools during their junior high years.

TABLE 2

Summary of Grades 8, 9 and 10 Standardised Mean Scores "F" ratios

	Grade 8	Grade 9	Grade 10	"F"
Creativity Measures	A 52.9	A 55.7	A 53.7	.87
	B 56.5	B 51.9	B 56.8	1.95
	C 51.5	C 55.7	C 58.0	2.63
Right Hemispheric Measures	A 49.5	A 58.1	A 60.0	7.35**
	B 52.8	B 55.0	B 58.0	1.54
	C 52.7	C 53.0	C 59.4	2.5
Left Hemispheric Measures	A 49.5	A 52.2	A 66.0	15.54**
	B 49.7	B 56.5	B 58.6	4.26*
	C 49.4	C 59.0	C 57.3	5.02**
Piagetian Development Measures	A 43.7	A 58.4	A 67.1	30.23**
	B 47.6	B 56.6	B 63.5	10.38**
	C 47.1	C 59.0	C 60.0	7.25**

** $p < .01$

* $p < .05$

Right Hemispheric Thinking vs. Grade.

The development of right hemispheric thinking abilities during this time seems to be less consistent. For school "A" right hemispheric abilities is related to the academic grade of the student but this is not true for schools "B" and "C". In examining the standardised mean score the variance in school "A" seems due to the difference between the grade 8 scores and scores of the two higher grades at the school. There also seems to be a large difference between the grades regards left hemispheric and Piaget thinking. Such a large difference could indicate that the grade 8 class is of lower ability in general terms. This suspected lower ability may have affected the lower right hemispheric scores.

Left Hemispheric Thinking vs. Grade.

The situation is very different with left hemispheric abilities which are significantly related to the grade of the student. This result reinforces the view that there is a development of left hemispheric thinking skills during the junior high years.

Piagetian Development vs. Grade.

Piagetian development is related to grade advancement according to this study. Such information again reinforces the developmental aspect of Piaget's theory.

CONCLUSION

In general terms the study indicates for the three schools examined the relationship between the academic grade of junior secondary advanced science students and creative right hemispheric thinking is weak, but the relationship between the student's grade and rational logical thinking is much stronger.

It may be that such relationships are a natural consequence of maturation and not related to the work done in the school. Most educators would feel that the school system has some effect.

If each child is going to realise one's full thinking potential, it is important that the creative and holistic modes of thinking be developed. In a future world which will be immersed in rapid change the need for this type of thinking is even more acute.

REFERENCES

ADAMS, L. Conceptual blockbusting. San Francisco: W.H. Freeman, 1974.

BARKER LUNN, J.C. Streaming in the primary school: Research report. London: National Foundation of Educational Research, 1970.

BURNEY, G.M. The construction and validation of an objective formal reasoning instrument. Unpublished Doctoral Dissertation. University of Northern Colorado, 1974.

BYRNE, P.S. & FREEMAN, J. Postgraduate training for general practice: An assessment of the aptitudes and abilities of trainee entrants. British Journal of Medical Education, 1971, 5(4), 292.

CROPLEY, A.J. & MASLANY, G.W. Reliability and factoral validity of the Wallach-Kogan creativity tests. British Journal of Educational Psychology, 1969, 40, 395-398.

DIMOND, S.J. & BEAUMONT, J.G. (Ed.) Hemisphere function in the human brain. New York: John Wiley & Sons, 1974.

FREEMAN, J., BUTCHER, H.J. & CHRISTIE, T. Creativity: a selective review of research, 2nd. Ed. London: Society for Research into Higher Education, 1971.

GETZELS, J.W. & JACKSON, P.W. Creativity and intelligence. New York: John Wiley & Sons, 1962.

GOODRUM, D. What is right about creative thinking? Science Education Conference, 1978, Western Australian Institute of Technology, Perth, 1978.

KOESTLER, A. The act of creation. New York: Macmillan, 1964.

MASLOW, A. Emotional blocks to creativity. In Parnes & Harding (Ed.) A source book for creative thinking. New York: Scribners, 1962.

ORNSTEIN, R.E. The psychology of consciousness. San Francisco: W.H. Freeman, 1972.

TISHER, R.P. A Piagetian questionnaire applied to pupils in a secondary school. Child Development, 1971, 42, 1633-1636.

TORRANCE, E.P. Educational achievement of the highly intelligent and highly creative: Eight partial replications of the Getzels-Jackson's study. Research Memorandum BER-60-18, Bureau of Educational Research, University of Minnesota, September, 1960.

WALLACH, M.A. & KOGAN, N. Modes of thinking in young children. New York: Holt, Rinehart & Winston, 1965.

THE CONTRIBUTORS

ALAN R. BARTON (DR)	Adelaide College of the Arts and Education
D.J. BOUD (DR)	Physics Department, Western Australian Institute of Technology
CHRIS DAWSON (DR)	University of Adelaide
JOHN DEKKERS (DR)	Science Education Centre, Western Australian Institute of Technology
ANNE DOWD	James Cook University of Northern Queensland
J.G. DUNN (DR)	Western Australian Institute of Technology
ROD FAWNS	Faculty of Education, Melbourne University
PETER FENSHAM (PROF)	Faculty of Education, Monash University
DARRELL L. FISHER	Tasmanian College of Advanced Education
JOHN GILBERT	Institute of Education Technology, University of Surrey
DENIS GOODRUM (DR)	Churchlands College of Advanced Education
LAWRENCE INGVARSON	Faculty of Education, Monash University
J.P. KEEVES (DR.)	Australian Council of Educational Research
NATALIE KELLETT (DR)	H.E.A.R.U., Monash University
T. KENNEDY	Western Australian Institute of Technology
ABDUL KHAN (DR)	Goulbern College of Advanced Education
RAMON LEWIS	Latrobe University
RUSSELL LINKE (DR)	School of Medicine, Flinders University
PADDY LYNCH	University of Tasmania
MAX MADDOCK (DR)	Faculty of Education, The University of Newcastle
JEFF NORTHFIELD	Faculty of Education, Monash University
ROGER OSBORNE (DR)	Department of Physics, University of Waikato
WIM PONNIAH (DR)	Faculty of Education, Monash University
COLIN POWER (PROF)	School of Education, Flinders University
LEONIE RENNIE	Department of Education, University of W.A.
GAYE ROSEN	University of N.S.W.
J. ROWELL	Department of Education, University of Adelaide

MICHAEL SULLIVAN	South Australian Department of Education
NOEL R. THORLEY	Western Australian Institute of Technology
M. VENZ	Flinders University

CONFERENCE COMMITTEE

MILES A. NELSON (Convenor)	Churchlands C.A.E.
MUREDACH DYNAN	Western Australian Institute of Technology
WILLIAM FOULDS	Claremont C.A.E.
PAT GARNETT	Nedlands C.A.E.
ROGER HACKER	University of Western Australia
GRAHAM PIKE	Mt. Lawley C.A.E.
RENATO A. SCHIBECI	Murdoch University